The Dissertation Committee for Daniel Ross Kagan
certifies that this is the approved version of the following dissertation:

Magnetic Instabilities and Resulting Energy Conversion in Astrophysics

Committee:

Milos Milosavljević, Supervisor

J. Craig Wheeler

Pawan Kumar

Anatoly Spitkovsky

John Scalo

Magnetic Instabilities and Resulting Energy Conversion in Astrophysics

by

Daniel Ross Kagan, B.S., M.A.

DISSERTATION

Presented to the Faculty of the Graduate School of
The University of Texas at Austin
in Partial Fulfillment
of the Requirements
for the Degree of

DOCTOR OF PHILOSOPHY

THE UNIVERSITY OF TEXAS AT AUSTIN
August 2013

Dedicated to my parents, Martin and Deborah Kagan.

Acknowledgments

I would first like to thank my supervisor, Milos Milosavljević; his guidance was critical for me, not only in guiding and collaborating with me in the portion of the dissertation he supervised, but also in giving me an understanding of how to do research and write as a scientist. I also thank J. Craig Wheeler, who collaborated with me on a major part of this dissertation, and was always there to talk to me about anything. I thank Anatoly Spitkovsky for providing the simulation code `Tristan-MP` and collaborating with me on the investigation of relativistic reconnection, and the other members of the `Tristan-MP` team at Princeton University for giving me excellent help with using the code. I also thank my other committee members, Pawan Kumar and John Scalo, and my former committee member Edward L. Robinson for their support and guidance. I thank Swadesh Mahajan for his collaboration and supervision on an early project and for supporting me in my later endeavours. Finally, I would like to thank my parents, Martin and Deborah Kagan, for their unwavering support throughout the long process of obtaining a Ph.D. And to all of those who after learning that my topic of study was astronomy immediately said "Oh, cool!": Thank you!

Magnetic Instabilities and Resulting Energy Conversion in Astrophysics

Publication No. __________________

Daniel Ross Kagan, Ph.D.
The University of Texas at Austin, 2013

Supervisor: Milos Milosavljević

Because the universe is primarily composed of plasma, the interaction of plasmas and magnetic fields is of great importance for astrophysics. In this dissertation, we investigate three magnetic instabilities and examine their possible effects on astrophysical objects. First, we model solar coronal structures as Double Beltrami states, which are the lowest energy equilibria of Hall magnetohydrodynamics. We find that these states can undergo a catastrophe with characteristics similar to those of a solar eruption, such as a flare or coronal mass ejection. We then investigate magnetic reconnection and particle acceleration in moderately magnetized relativistic pair plasmas with three-dimensional particle-in-cell simulations of a kinetic-scale current sheet. We find that in three dimensions the tearing instability produces a network of interconnected and interacting magnetic flux ropes. In its nonlinear evolution, the current sheet evolves toward a three-dimensional, disordered state in which the resulting flux rope segments contain magnetic substructure on kinetic scales and sites of temporally and spatially intermittent dissipation. We find that reconnection produces significant particle acceleration, primarily due to the electric field in the X-line regions between flux ropes; the resulting particle energy spectrum can extend to high Lorentz factors. We find that the highest energy particles are moderately beamed within $\sim 30^\circ - 40^\circ$ of the direction of acceleration. Finally, we derive a dispersion relation and calculate growth rates for triply-diffusive nonaxisymmetric instabilities including the magnetorotational instability (MRI) throughout the Sun, accounting for the effects of both shear and convective buoyancy. The overall instability has unstable modes throughout the convection zone and at colatitudes $\theta < 53^\text{o}$ in the tachocline. The instability contains three classes of modes: large-scale hydrodynamic convective modes, large-scale hydrodynamic shear modes, and small-scale MRI shear modes. While large-scale convective modes are the fastest-growing modes in most of the convective zone, MRI modes are important in both stably stratified and convectively unstable locations near the tachocline at colatitudes $\theta < 53^\text{o}$. We find that nonaxisymmetric MRI modes typically grow faster than axisymmetric MRI modes. We consider the saturation of magnetic fields produced by the MRI, finding that they may be comparable to those produced by a convective dynamo.

Table of Contents

List of Tables

List of Figures

Chapter 1

Introduction

Because most of the universe consists of ionized plasma that can grow and sustain magnetic fields, these fields play a critical role in many astrophysical systems, from the Earth's magnetosphere up to galactic scales. The initial growth of magnetic field in the universe may be a result of the Biermann battery dynamo mechanism (Widrow et al., 2012). Once a seed galactic magnetic field is present, it can can grow through the interaction of shear and convection. The resulting field in the interstellar medium plays an important role, via magnetic pressure, in determining which areas of a gas cloud will collapse to form stars. The collapse of a subcloud results in the formation of a protostar surrounded by a protoplanetary disk; the rapid growth of the magnetorotational instability (MRI) can then provide an effective viscosity that transports angular momentum in the disk and allows the accretion of material onto the star, while significantly amplifying the magnetic field (Balbus & Hawley, 1998, and references therein). In the inner disk, accreting material is diverted by the magnetic fields of the star and the accretion disk to produce the stellar winds and jets observed near the polar axes of T Tauri protostars (McKee & Ostriker, 2007, and references therein). During a star's main sequence lifetime, its magnetic field may play a role in regulating the differential rotation of the star. In addition, the buildup of magnetic energy in the stellar corona and its release through magnetic reconnection or the global loss of equilibrium is responsible for a large portion of stellar activity; see Balogh & Erdõs (2013) for a review of these processes in the Sun. In our own solar system, the coronal mass ejections produced by the Sun's activity often reach the Earth, where they can do significant damage to satellites; they would do significantly more harm if not for the shielding provided by Earth's own magnetic field, produced by its rotating iron core.

Magnetic fields also play an important role in the astrophysics of neutron stars and black holes. Neutron stars can directly undergo spin-down by means of their dipole magnetic field. Many black holes and neutron stars of these objects are observable only when surrounded by accretion disks. Magnetic processes similar to those in protostellar disks can result in accretion mediated by the MRI and the launch of a magnetized jet. For compact objects, this jet is typically relativistic, especially in the case of black holes.. The larger-scale black holes that are responsible for active galactic nuclei (AGN) result in jets that can have a strong effect on galaxy formation and evolution (Fabian, 2012, and references therein). Gamma-ray bursts of both long and short duration are thought to be associated with the formation of a black hole fed at an extremely high rate by the fast accretion of a disk and the launch of an ultrarelativistic jet (Piran, 2004, and references therein).

The observed radiation resulting from solar activity, supernova remnants, and relativistic jets is typically highly energetic and nonthermal, and magnetic fields are deeply involved in the production of this radiation. The most promising processes for accelerating particles to high energies are collisionless Fermi acceleration (Bell, 1978; Blandford & Ostriker, 1978), in which a particle repeatedly bounces between magnetic domains on either side of a shock, and magnetic reconnection (e.g. Zenitani & Hoshino, 2001), in which a particle is accelerated directly by the electric field as magnetic energy is converted to kinetic energy; however, the efficiency of these mechanisms in specific astrophysical systems is not well understood. If these mechanisms are effective, they produce a significant population of high-energy charged particles; the most strongly accelerated of these particles may reach the Earth and be observed as cosmic rays. The typical radiation mechanism for these high-energy particles is thought to be synchrotron radiation resulting from gyrations about magnetic field lines, possibly boosted by inverse Compton scattering. Thus, the magnetic field plays an integral role in both the evolution of astrophysical systems and the production of observed emission.

1.1 Models of Plasmas

In general, an astrophysical plasma must be described by specifying the electric field **E** and magnetic field **B** at each point in space, and the position **x** and momentum **p** of each particle in the plasma. In the classical limit, evolution of the system is determined in general by Maxwell's Equations and the equation of motion for each particle, which includes both the electromagnetic Lorentz force and other forces, such as a gravitational force or binary collisions[1]. In many astrophysical situations, however, collisions are infrequent and the effects of gravity may be

---

[1] Binary "collisions" are also the result of the electromagnetic force, but the extremely short scale of the fields responsible for the collision means that it is more common to treat collisions separately.

neglected. Then, the equation of motion of each particle is

$$\frac{d\mathbf{p}}{dt} = q\left(\mathbf{E} + \frac{1}{c}\mathbf{v}\times B\right). \tag{1.1}$$

where **v** is the velocity of the particle and q is its charge. In Chapter 3 of this dissertation, we use Maxwell's Equations calculated on a grid combined with Equation (1.1) to simulate the evolution of an ensemble of macroparticles that each represent many charged particles; this approach is capable of treating all effects present in plasma physics and allows the detection of particle acceleration to high energies.

The equations of motion of the particles may be simplified considerably using the phase space density $f_s(\mathbf{x},\mathbf{p})$ of each species of particles in position and momentum space. In the absence of collisions, the phase space density will be conserved in the Lagrangian sense, resulting in the Vlasov equation of motion for each species s

$$\frac{\partial f_s}{\partial t} + (\mathbf{v}_s\cdot\nabla)f_s + q_s\left(\mathbf{E} + \frac{1}{c}\mathbf{v}_s\times\mathbf{B}\right)\cdot\frac{\partial f_s}{\partial \mathbf{p}_s} = 0. \tag{1.2}$$

Combining this equation with Maxwell's equations again specifies the complete evolution of the system.

1.1.1 Kinetic Effects

The use of the Vlasov Equation along with Maxwell's Equations is difficult because the Vlasov Equation requires specification in six dimensions, while Maxwell's Equations use only three dimensions and require the calculation of the local charge density and current from the phase space density f. However, the presence of electromagnetic fields typically means that the motion of particles on small scales is not isotropic unless collisions are very frequent. Because the magnetic force on a particle is perpendicular to the direction of the particle's motion, a particle's motion in the plane perpendicular to the magnetic field will describe a circle with radius r_{L} given by

$$r_{\mathrm{L}} = \frac{p_{s,\perp}}{q_s B}. \tag{1.3}$$

where $p_{s,\perp}$ is the magnitude of the particle's momentum perpendicular to the magnetic field. Motion in the parallel and perpendicular directions is therefore significantly different on length scales smaller than r_{L}.

The other important kinetic scale is the frequency of plasma oscillations $\omega_{\mathrm{p},s}$. The corresponding kinetic length scale is the skin depth $\lambda_{\mathrm{p},s} = c/\omega_{\mathrm{p},s}$, which is given by

$$\lambda_{\mathrm{p},s} = \sqrt{\frac{\langle U_s \rangle}{4\pi n_s q_s^2}}, \tag{1.4}$$

where $\langle U_s \rangle$ is the average energy of particles of species s and n_s is the number density of particles of that species. On length scales significantly smaller than $\lambda_{\mathrm{p},s}$, electric fields can penetrate into a plasma even in the absence of collisional resistivity.

In general, the inclusion in a plasma model of physics on these small scales requires a full kinetic treatment using the Vlasov Equation or separate equations of motion for each particle because the particle momentum distributions cannot be treated as thermal on small scales. However, when these small-scale effects are unimportant, further simplifications can be made.

1.1.2 Flux Freezing and Fluid Models

If small-scale kinetic effects can be neglected, it is reasonable to assume that the particle momentum distribution is thermal. The combination of these assumptions yields the simplest fluid model of plasma physics, magnetohydrodynamics (MHD). In MHD, the plasma can be treated as a single fluid with center-of-mass fluid velocity $\langle \mathbf{v} \rangle$ and current density $\mathbf{J}$. In this theory, the electric field E is related to the magnetic field $\mathbf{B}$ and average velocity $\langle \mathbf{v} \rangle$) by the equation:

$$\mathbf{E} + \frac{1}{c} \langle \mathbf{v} \rangle \times \mathbf{B} = 0 \tag{1.5}$$

This equation implies that the fluid is frozen onto magnetic field lines, so that it cannot move perpendicular to them. MHD can be easily modified to include collisional resistivity η by adding a term of the form $\eta \nabla \times \mathbf{B}$ to the right side of Equation (1.5), producing resistive MHD. Simple diffusive effects such as viscosity and thermal conductivity can be added by modifying other parts of the MHD equations in similar fashion. In Chapter 4, we use MHD including diffusive effects to calculate the dispersion relation and growth rates of the magnetorotational instability in the Sun.

In electron-ion plasmas, some kinetic effects can be included in a fluid theory by making use of the fact that the Larmor radius and skin depth of ions are much larger than those of the electrons, due to their large differences in mass. Therefore, on scales between the ion and electron kinetic scales one can replace the flux-freezing approximation Equation (1.5) with the assumption that the flux is frozen into the *electron* fluid

$$\mathbf{E} + \frac{1}{c} \langle \mathbf{v}_e \rangle \times \mathbf{B} = 0. \tag{1.6}$$

The resulting fluid theory is called Hall MHD; the fact that $\langle \mathbf{v}_e \rangle$ enters into the equations means that the center-of-mass velocity field $\langle \mathbf{v} \rangle$ has dynamical effects, rather than being frozen into the magnetic field. In Chapter 2 of this dissertation, we use Hall MHD to model coronal structures that have steady velocity flows.

1.2 Instabilities Discussed in this Dissertation

In this dissertation, we investigate several instabilities in plasmas that are of importance in astrophysics. We now introduce them, and explain their importance in astrophysics.

1.2.1 Loss of Equilibrium

In Chapter 2 of this dissertation, we find that solar coronal structures can be modelled as Double Beltrami states, which are the lowest-energy equilibria of Hall MHD. Such equilibria allow for the presence of small-scale fields and steady velocity flows in equilibrium. When coronal structures evolve slowly, one can model their evolution as a progression between adjacent Double Beltrami equilibria. However, we find that in some circumstances, there may be no adjacent equilibrium to which to move. In this case, a loss of equilibrium, or catastrophe, occurs. This implies that some nonideal and discontinuous process is required to reach a final equilibrium state. This transition is accompanied by the rapid conversion of magnetic energy to kinetic energy of the flow. It is uncertain what process will ensue when the equilibrium is lost, but observations of solar eruptive events indicate that they are often associated with magnetic reconnection (Forbes et al., 2006), which is produced by the tearing instability.

1.2.2 The Tearing Instability and Reconnection

Whenever magnetic domains of opposite polarity are present, Maxwell's equations imply that there will be a current sheet, or a region of high current, in between them. The oppositely—oriented fields constitute a source of free energy, since the fields can cancel eachother if plasma from both sides of the current sheet moves into the current sheet. The primary instability of a current sheet that draws on this energy is the tearing instability, which produces an alternating series of narrow X-lines and wide flux ropes. Plasma flows into the X-points from outside the current sheet in the magnetic domains, and then flows outward, parallel to the plane of the the current sheet, into the flux ropes. At the center of the X-points, magnetic fields of opposite polarity reconnect and are ejected into the flux ropes. In turn, steady reconnection equilibria contain thin current sheets, which themselves can be unstable to the tearing instability. This indicates that there is a strong connection between the tearing instability and magnetic reconnection.

Because reconnection of a field line at the X-line violates flux freezing for the plasma which threads that field line, the tearing instability requires the presence of some nonideal effect not present in MHD. The nonideal effect that violates flux freezing may be collisional resistivity, or it may arise from kinetic effects. Therefore, the tearing instability is both a resistive MHD instability and a kinetic instability. The nonideal electric field present at the center of the X-line thet converts significant amounts of magnetic energy into kinetic energy and it may be able to accelerate particles to high energy (e.g. Zenitani & Hoshino, 2001). The growth rate of the tearing instability depends strongly on the width of the current sheet; for fast growth, the sheet width must be comparable to resistive or kinetic length scales. In Chapter 3 of this dissertation, we carry out three dimensional simulations of the kinetic tearing instability in a relativistic current sheet with width close to the plasma skin depth and investigate energy conversion and the resulting particle acceleration.

1.2.3 The Magnetorotational Instability

In both accretion disks and stars, differential rotation can represent a source of free energy that drives instability. In the hydrodynamic case, the differential rotation leads to instability only if the specific angular momentum decreases outward (Tassoul, 1978). However, the presence of even a small magnetic field in a plasma has a drastic effect on the stability condition; instability occurs as long as the angular *velocity* decreases outward, a condition met in all accretion disks and in some cases in stellar interiors and during core collapse. The resulting MHD instability is called the magnetorotational instability (MRI), was independently discovered by Velikhov (1950) and Chandrasekhar (1960), but its significance in astrophysical contexts was first recognized by Balbus & Hawley (1991). A useful way of understanding the physical origin of this instability is presented by Balbus & Hawley (1998): consider two masses in adjacent orbits connected by a spring which represents the magnetic field that is frozen into the plasma represented by the masses. The mass in the lower orbit is pulled backwards by the spring, since the increase in angular velocity with decreasing radius implies it moves forward faster than the mass in the higher orbit; therefore, it loses angular momentum and moves to a lower orbit, and its speed increases. The opposite is true for the mass in the higher orbit: it slows down and moves into an even higher orbit. Thus, the difference in angular momentum between the masses increases with time, as does the tension in the spring, which corresponds to the strength of the magnetic field. The MRI typically has a very fast linear growth rate similar to the local angular velocity in systems like accretion disks in which the shear is strong; in its nonlinear development, this instability can produce strong turbulence that is now thought to be responsible for the effective viscosity that produces angular momentum transport in accretion disks (Balbus & Hawley, 1998). In Chapter 4 of this dissertation, we investigate the applicability of the MRI to the Sun, deriving a dispersion relation including buoyant, magnetorotational, and diffusive effects and calculating growth rates of both axisymmetric and nonaxisymmetric modes.

Chapter 2

The Double Beltrami Model of Solar Eruptions[1]

2.1 Introduction

Sudden and catastrophic events, such as solar flares, CMEs, and erupting prominences are common in the solar corona. These events probably occur too rapidly for energy injection from the photosphere to drive them directly (Krall et al., 2000); they may instead occur when energy, stored over time in the coronal magnetic field, is rapidly released. Despite years of study, there is still no consensus as to what mechanism is behind these eruptive events.

One method for modelling solar eruptions is to assume an initial ideal magnetohydrodynamic (MHD) equilibrium and examine if a breakdown of the equilibrium can harness enough energy to produce an eruptive event. Many such equilibria are not energetically favorable; small perturbations, then, can drive the system to a new relaxed, lower energy equilibrium. The only equilibrium that is truly stable is the most likely equilibrium, found by minimization of the energy subject to topological constraints.

In magnetohydrodynamics (MHD), the simplest fluid model for plasma physics, it can be shown that the minimum energy static (V=0) equilibrium is the

---

[1] A large part of the material in this chapter is taken from the manuscript Kagan, D., Mahajan, S. M., 2010, Mon. Not. Roy. Ast. Soc., 406, 1140.

force-free magnetic field (Woltjer, 1958):

$$\nabla \times \mathbf{B} = \lambda \mathbf{B}, \tag{2.1}$$

where λ is a constant determined by the the ratio of energy to the magnetic helicity, the latter being the topological invariant of ideal MHD. Fitting a force-free field to magnetic field measurements is a commonly used practice in modeling coronal structures.

While the force-free model of MHD is excellent at modeling equilibrium structures, it cannot account for the breakdown of these equilibria associated with changes (via reconnection, for instance) in the field topology. Such a change must occur, because an open (potential) magnetic field that results from a reconnection event has a higher magnetic energy than any corresponding initial closed field (Aly, 1991; Sturrock, 1991); no ideal MHD process can drive such an eruption.

Most models assume that eruptive events are driven by the reconnection of magnetic fields due to non-ideal effects like resistivity; an appropriate field geometry that can produce a reconnection event is assumed. In two standard models, initial equilibrium is assumed to be consisting, respectively, of flux ropes and multi polar magnetic fields; the latter situation defines the so called breakout model (Forbes et al., 2006); the pre-eruption coronal structures, however, may have neither of these configurations. In addition, collisional resistivity is not large enough in the corona to produce fast reconnection. Therefore, MHD models of solar eruptions require the presence of ‘anomalous’ resistivity and special assumptions about the magnetic field topology in the corona.

Because of high conductivity, the Hall effect is more important than resistivity in the solar corona. It can lead to fast reconnection by providing small-scale fields that dissipate more quickly (Bhattacharjee, 2004); it can also lead to the formation of fast flows from turbulent magnetic fields using the reverse dynamo mechanism (Mahajan et al., 2005), and the formation of coronal structures from an initial upflow into the corona, including the relatively fast flows observed in coronal loops (Mahajan et al., 2001). These successes indicate that Hall MHD may allow the construction of a successful "‘loss of equilibrium"’ model for catastrophic events without making the special assumptions about topology as the flux rope and breakout models do.

This paper, investigating the possible role of Hall effect in solar eruptions, is organized as follows. In Section 2.2, we first describe the most likely equilibrium accessible to Hall MHD, the Double Beltrami (DB) State, and then show that DB states can suffer catastrophic loss of equilibrium under solar coronal conditions. In Section 2.3, we compare the characteristics of the DB catastrophe with those of solar eruptions, and compare the predicted expansion during the catastrophe with observed CME kinematics. In Section 2.4, we calculate the rate of CMEs resulting from the DB model using a phase space argument, and compare it with the observed rate. Finally, in Section 2.5, we summarize our conclusions and discuss possible future work.

2.2 Hall MHD and Double Beltrami States

The equations of Hall MHD are

$$\frac{\partial \mathbf{B}}{\partial t} = \nabla \times ((\mathbf{V}-\mathbf{J}) \times \mathbf{B}), \tag{2.2}$$

$$\frac{\partial \mathbf{V}}{\partial t} = \mathbf{V} \times (\nabla \times \mathbf{V}) + \mathbf{J} \times \mathbf{B} - \nabla(P + \Phi_g + \frac{V^2}{2}). \tag{2.3}$$

In these equations **J** is the current density, P is the pressure, and Φ_g is the gravitational potential. We use units in which **B** is normalized to an appropriate B_0, **V** is normalized to the corresponding Alfven velocity, and the length scale is the size of the system (this sets the value of the time scale). The equivalence of electrons and ions that exists in ordinary MHD, which is what allows the dynamics to be described using only the evolution of the ions, is broken in Hall MHD; the magnetic field is determined by both the electron velocity $\mathbf{V}_e$ and the ion velocity $\mathbf{V}_i$, since $\mathbf{J} = \mathbf{V}_i - \mathbf{V}_e$, while $\mathbf{V} \approx \mathbf{V}_i$. The result is that in equilibrium the electrons still follow the magnetic field lines as before (since $V_e = V - J$, the electrons' induction equation is identical to that of ordinary MHD), but the ions instead follow the field lines minus the vorticity of the field lines (the vorticity of a field is its amount of circulation, or its curl). By setting the time derivatives of these equations to zero, we can find equilibria of these equations, and then compare the equilibria to observations.

The most likely equilibrium in Hall MHD results from the minimization of energy E while the magnetic helicity (electron helicity) h_e and the generalized helicity (ion helicity) h_i are held constant. The resulting equilibrium Double Beltrami (DB) state has magnetic field **B** and velocity field **V** that consist of a sum of two simple Beltrami fields on two different scales (Mahajan & Yoshida, 1998):

$$\mathbf{B} = C_{\mathrm{L}}G_{\mathrm{L}} + C_{\mathrm{S}}\mathbf{G}_{\mathrm{S}}, \tag{2.4}$$

$$\mathbf{V} = (\lambda + \widetilde{a})C_{\mathrm{L}}\mathbf{G}_{\mathrm{L}} + (\mu + \widetilde{a})C_{\mathrm{S}}\mathbf{G}_{\mathrm{S}}. \tag{2.5}$$

where the **G** fields satisfy the Beltrami equation (1). Note that **B** and **V** are dimensionless, as they have been normalized to some magnetic field B_0 and the corresponding Alfven velocity. The constants $C_{\mathrm{L,S}}$ are the amplitudes of the large and small scale fields that have characteristic inverse scale lengths λ and μ ($\widetilde{a}$ is a parameter that may be calculated from these scale lengths).

2.2.1 Breakdown of double Beltrami equilibria

The Double Beltrami equilibrium has more energy available to drive an eruptive event than the linear force-free state. It can undergo a catastrophe converting magnetic energy into kinetic energy, thus, simulating one of the defining features of a eruptive event. Ohsaki, Shatashvilli, Yoshida, and Mahajan (2002) (henceforth, OSYM) found that a sequence of slowly varying Double Beltrami equilibria constrained by the invariants can, indeed, terminate in a catastrophe; at the catastrophe boundary the amplitude of the small scale vanishes tending to (formally) becoming imaginary. OSYM found that the removal of the small scale can occur if the system energy exceeds a critical energy determined by the two helicities:

$$E > E_{\mathrm{c}} \equiv 2(\sqrt{h_e h_i} - h_e). \tag{2.6}$$

Since the DB state can undergo a catastrophy, it is a possible state from which a solar eruption could occur. However, in order to construct a model of solar eruptive events, we must identify an evolutionary path for the four parameters that define a DB state in which the initial state has properties that correspond to those of coronal structures and the final state has the defining properties of an eruptive event.

2.2.2 Observational constraints on the initial state

To use the DB model to simulate observations, we normalize the Double Beltrami states as follows: We normalize the magnetic field **B** to the macroscopic field of an observed coronal structure (this sets $C_{\rm L} = 1$), and the velocity **V** to the corresponding Alfven speed. The normalizing speed is chosen to be the Alfven speed based on the large-scale field alone; this is very close to the true Alfven speed in our parameter region of interest.

Observations of the solar corona constrain the energy E, large scale wavenumber λ, and large scale normalized velocity $v \equiv V_{\rm L}/v_{\rm A}$ of the initial state. The maximum flow velocities commonly seen in loops are generally about 100 km s^{-1} (Kjeldseth-Moe & Brekke, 1998; Fredvik et al., 2002), while the Alfven speed for flux ropes in the solar corona (which have $B \approx 10g$, $n \approx 10^8 cm^{-3}$) is approximately 2000 km s^{-1}. This requires that $v \approx .01 \ll 1$. Now we look at the energetics of a typical eruptive event. A catastrophe can only occur with $v \ll 1$ if $E-1 \gg 1$ or $E-1 \ll 1$; ie, if almost all of the energy is in the macroscopic magnetic field, or almost all is outside it. The energy of solar flares is consistent with a pre-catastrophe situation in which the loop energy lies primarily in the macroscopic magnetic field. This restricts the realistic region for the initial state to $(E-1 \ll 1, v \ll 1)$ where $\lambda \ll 1$, $\mu \gg 1$, and $C_{\rm S} \ll 1$.

We now calculate the fields and invariants for the region of interest; the expressions are simplified greatly by assuming $\lambda \ll \mu^{-1}$; this implies that $\widetilde{a} \approx 1/\mu \approx v$ for typical coronal structures (OSYM). The size of typical structures in the corona is approximately 1 Mm (ie, about one solar radius), while the Hall scale is typically about 1 m in the corona. This allows us to estimate $\lambda \approx 10^{-6}$ for these structures. Flows in the corona have typical velocities of tens of km s^{-1}, while the Alfven speed in the Corona is of order 1000 km s^{-1}; this sets $v \approx \mu^{-1} \approx .01$. So our assumption that $\lambda \ll \mu^{-1}$ is justified for typical coronal structures.

The invariants and fields are

$$E \approx 1, \tag{2.7}$$

$$h_e \approx 1/\lambda, \tag{2.8}$$

$$h_i - h_e \approx C_{\mathrm{S}}^2 \mu^3, \tag{2.9}$$

$$\mathbf{B} \approx \mathbf{G}_{\mathrm{L}}, \tag{2.10}$$

$$\mathbf{V} \approx \mu^{-1} \mathbf{G}_{\mathrm{L}} + \mu C_{\mathrm{S}} \mathbf{G}_{\mathrm{S}}. \tag{2.11}$$

The magnetic field is mostly large-scale, and close to being force-free. The velocity field is small in magnitude but exists on both large and small scales, and is consistent with the presence of observed slow flows on large scales as well as small-scale turbulence.

Because $h_i - h_e \ll h_i$, the critical energy may be approximated as

$$E_{\mathrm{c}} = 2(h_e - \sqrt{h_e h_i}) \approx h_i - h_e \approx C_{\mathrm{S}}^2 \mu^3. \tag{2.12}$$

A catastrophe can occur if $C_\mathrm{S}^2\mu^3 < 1$.

2.2.3 Energy produced by the catastrophe

As the system approaches the catastrophe while conserving the invariants, the macroscopic magnetic field and velocity change, as do the amplitudes C_L and C_S and the inverse scale lengths λ and μ. At the point of catastrophe, $C_\mathrm{S} = 0$, and the field is entirely macroscopic. We now test whether the catastrophe produces enough energy to drive an eruptive event. We will use uppercase greek letters for parameters at the catastrophe. At the critical point, the large scale Λ is given by (OYSM)

$$\Lambda = \frac{E+\sqrt{E^2-E_\mathrm{c}^2}}{2h_e}. \tag{2.13}$$

Using this equation, it is then straightforward to calculate the macroscopic velocity at the catastrophe (in Alfven units)

$$U^2 = \frac{E-\sqrt{E^2-E_\mathrm{c}^2}}{2}. \tag{2.14}$$

The maximum velocity producible by the DB model occurs at $E_\mathrm{c} = 1$, and with $E \approx 1$, it is $U = 1/\sqrt{2}$. At this velocity, half of the magnetic energy has transformed into kinetic energy.

2.3 Comparison with Observations

In this section we attempt to justify the model by comparing the properties of the initial and final DB states, respectively, with the initial conditions in the corona and final conditions observed in a solar eruptive event.

2.3.1 Eruptive velocities

First we must make a comparison of the observed velocities of eruptive events with with those predicted at the catastrophe. Flux ropes in active regions in the solar corona have Alfven speeds of approximately 2000 km s^{-1} (Chen, 2001); the DB catastrophe can produce a final velocity of $v_\mathrm{A}/\sqrt{2} \approx 1400$ km s^{-1}. The vast majority of CMEs have peak velocities less than 1500 km s^{-1} (Schwenn et al., 2006). Therefore, the most energetic DB catastrophes should be able to produce enough energy to create CMEs; less energetic DB catastrophes can produce smaller flares or prominence eruptions.

2.3.2 Changes in the DB parameters

Near the catastrophe, $h_e \approx 1/\lambda$, $E \approx 1$, and λ behaves according to the equation

$$\frac{\Lambda}{\lambda} \approx \frac{1+\sqrt{1-E_\mathrm{c}^2/E^2}}{2}. \tag{2.15}$$

The minimum ratio between the final and initial large scale size is 1/2, at $E = E_\mathrm{c}$; for $E > E_\mathrm{c}$, the ratio will be closer to 1. So λ decreases slightly as the system moves towards the critical point; therefore the large scale increases in size by a small amount. The change from the initial to the final length scales corresponds well with known properties of coronal loops (initial), in which the sides tend to repel due to the Shaforonov hoop stress, and the properties of CMEs and prominences (final) that expand and move upward as they approach the moment of eruption.

Next, we consider the value of C_L (which is also the amplitude of B_L). For the region of interest, $h_e \approx C_\mathrm{L}^2/\lambda$. Since h_e is invariant, $C_\mathrm{L} \propto \sqrt{\lambda}$. The large scale magnetic field, thus, drops as the system approaches the catastrophe. The decrease in the magnitude of the macroscopic magnetic field fits well with what we see in eruptive events; magnetic energy is converted into kinetic energy or radiation.

Next, we consider the changes in the small scale parameters. At the catastrophe point, the small scale amplitude is 0, because it is there that the Double Beltrami state is converted to a purely macroscopic state. The small scale length $1/\mu$ grows to a value close to the Hall scale (from an initial size about 100 times smaller). It is possible that this small scale size corresponds to the width of a current sheet produced in the DB catastrophe; most models of solar reconnection assume that fast reconnection is mediated by the presence of a small-scale field at approximately the Hall scale. Because neither the Double Beltrami state nor the large-scale final field are exactly force-free due to the presence of a nonzero flow velocity), the DB model can accommodate the presence of current sheets. However, it is unnecessary to invoke a current sheet to produce an eruptive event in the DB catastrophe, because it transforms magnetic energy to kinetic energy without assumed dissipation.

2.3.3 Result of the catastrophe

While the general features of the post-catastrophe state discussed previously can be predicted using the DB equilibrium model, detailed features of this state (the detailed field geometry, the fate of the kinetic energy) cannot. The region modelled by the DB state may erupt off the solar surface (if its velocity exceeds the solar escape speed); this may correspond to a CME. It is also possible that the energetic plasma will be confined by the local magnetic field; this is especially likely to occur if the eruptive event has low energy compared to that in the local magnetic field ($U \ll 1$). If this occurs, most of the kinetic energy is likely to be converted to heat and radiated away, resulting in a non-eruptive flare.

Because the large-scale velocity field varies within the DB region, the results of the catastrophe may vary from place to place; a CME and a flare could result from the same catastrophe. This variation may also result in the formation of small-scale structures such as current sheets. Investigation of these questions is beyond the scope of this paper. However, it is possible to make two quantitative comparisons between the DB model and CMEs; these tests will serve as indicators of the applicability of DB states to solar eruptions in general.

2.3.4 CME kinematics

The DB model predicts that the scale size of the state will increase by a factor of 2 or less before the catastrophe occurs; furthermore, the higher the final energy of the eruptive event, (in Alfvenic units) the greater the expansion. The kinematics of CMEs, which often display a quasi-equilibrium stage prior to eruption, provides an excellent observational test for this pre-catastrophe expansion.

2.4 Calculating the rate of CMEs

Most CMEs display a three-part kinematic history (Zhang et al., 2001). The first may be thought of as a quasi-equilibrium stage during which structures rise slowly, with little change in speed. The second phase is associated with rapid acceleration; it is in this stage that the CME nearly reaches its final velocity. Finally,

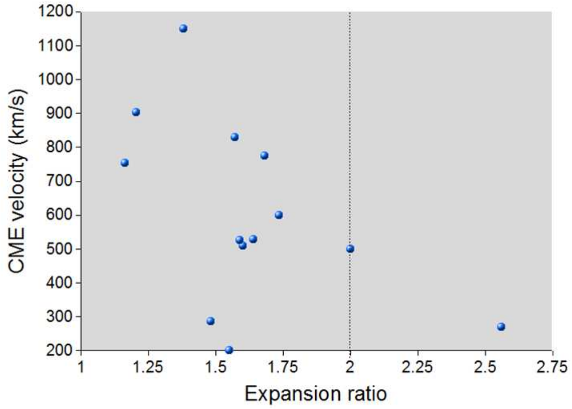

Figure 2.1 A comparison of the quasi-equilibrium expansion rate of CMEs to their final velocity. The CME data was taken from the LASCO C1 database.

the CME drifts at near-constant speed, affected mostly by drag forces.

Because the first stage corresponds well with the quasi-equilibrium model, measuring the height at which acceleration begins in comparison with the height at which the leading edge of the CME is first viewed, and then finding the ratio of the heights, will be a the test of the theory. We use CME events from the LASCO C1 CME database (which can be found at `http://solar.scs.gmu.edu/research/cme_c1`) to test the prediction. Only CMEs with a clear quasi-equilibrium stage and a clearly visible leading edge are included in the sample.

The figure indicates that most of the CMEs meet the criterion of a two fold or smaller expansion, and the others do not exceed the criterion by much. Even these anomalous CMEs are not necessarily in contradiction with the model, because the height of the leading edge is only one dimension of the coronal field; the others could expand less, leading to a smaller average value of λ. We do not observe the expected correlation (2.15) between the CME speed and the amount of expansion. This discrepancy may be due to the variation in the Alfven speeds at the locations of the CMEs; the predicted relationship is between *normalized* velocity and expansion ratio. CME kinematics are therefore consistent with the model, although more observations (with approximate measures of the Alfven speed) are needed.

If DB states that form in the solar corona are chosen randomly from possible coronal states, and eruptions are independent events, the rate at which a catastrophic eruption occurs in the solar corona R_{C} can be calculated from the rate R_{D} at which new coronal states form:

$$R_{\mathrm{C}} = \frac{A_{\mathrm{C}}}{A_{\mathrm{D}}} R_{\mathrm{D}}. \tag{2.16}$$

Here A_{C} is the phase space area of catastrophe-prone realistic DB states and A_{D} is the phase space area of all realistic DB states. Because the simplest result of a DB catastrophe is a CME, and CMEs are the most violent and visible solar eruptive events, we choose to estimate the CME rate rather than that for some other eruptive event.

2.4.1 The predecessor and its formation rate

To calculate the predicted rate of CMEs, it is necessary to specify a structure on the Sun that corresponds to the DB state; the CME is then hypothesized to be the result of the breakdown of equilibrium that occurs if the structure is catastrophe-prone. In doing so, we assume that all catastrophe-prone states eventually undergo catastrophe, and that the formation and ejection rate of catastrophe-prone states is approximately the same.

About 85% of CMEs are associated with active regions (Dasso et al., 2005) (the rest may be associated with ones that recently appeared or dissapeared), and CMEs are large-scale events that cover entire active regions; therefore, it makes sense to identify DB states with active region configurations. The timescale for large-scale changes in an active region is generally on the order of days, but it can be as short as 1 hour during times of heightened activity, when flares or CMEs occur. It therefore makes sense to assume that the DB state in the active region, associated with its large-scale structure, is continually modified by emerging coronal loops; the result is that a new DB state is formed every hour in an active region, and the emergence (and removal) of flux can move the state towards or away from a CME-prone state. The small “distance" that a catastrophe-prone state must travel in phase space to reach a loss of equilibrium means that the catastrophe is likely to occur quickly, before a new DB state replaces the old one.

We can estimate upper and lower limits on the timescale on which the DB state is replenished. A maximum timescale for the formation of CME precursors may be found by examining successive CMEs from the same region; the precursor formation rate for active regions must be greater than the rate at which successive CMEs can occur in the regions of greatest activity. The recurrence time for CMEs in NOAA active region 9236 was found to be about 4.6 hours, while the smallest time difference was shorter (Gopalswamy et al., 2006). A minimum timescale may be found by calculating the radiative cooling time for typical coronal loops; any major change cannot happen faster than the time required for a loop to destabilize, about 10 minutes (Kjeldseth-Moe & Brekke, 1998). So a 1 hour estimate for DB state emergence in an active region seems reasonable.

To allow for more accurate results with less net variability, all parameters will be calculated at solar maximum,when there are typically 10 active regions on the Sun (this was obtained from `http://www.solarmonitor.org`) by sampling timesteps near the maximum). In addition, the sunspot number on the Sun is about 100-150 at solar maximum, which corresponds to 10 active regions if the number of solitary sunspots is low. On the whole sun, then, the resulting rate of production of CME predecessors is 240 d^{-1}.

2.4.2 Phase space areas

The projection of the phase space of DB states that is relevant to the probability of catastrophe is 2-dimensional, because the probability of catastrophe and the CME velocity provided by the catastrophe depend only on E_C/E, and for the regime of interest, $E_C/E \approx C_S^2\mu^3$. Therefore, we must specify the ranges of C_S and μ that fall into A_C and A_D.

The value of C_S cannot be constrained by any known observation. In contrast, the value of the small scale $\lambda_S \approx 1/v_0$, since all predecessor velocities are very sub-Alfvenic; therefore, constraints are possible due to limitations on the initial velocity. For the initial state, the boundaries of the region are marked by just the values of μ corresponding to the maximum and minimum of v (the velocity of flows in the predecessor), and the unknown bounds of C_S. For the final state, the bounds are

$$2U_{\mathrm{h}}\sqrt{1-U_{\mathrm{h}}^2} > C_{\mathrm{S}}^2\mu^3 > 2U_{\mathrm{l}}\sqrt{1-U_{\mathrm{l}}^2}, \tag{2.17}$$

where U_{h} is the maximum critical velocity that corresponds to a CME, and U_{l} is the minimum critical velocity.

The initial state boundaries are just the boundaries in $v \approx 1/\mu$, while the ratio of $E_{\mathrm{C}}/E \approx C_{\mathrm{S}}^2\mu^3$ values is the same for all C_{S}. Therefore, a calculation of the logarithmic areas eliminates C_{S} and allows a prediction of the CME production rate. The resulting estimates of the phase space areas are

$$A_{\mathrm{C}} = (\log\frac{C_{\mathrm{S,h}}}{C_{\mathrm{S,l}}})((\log(\frac{U_{\mathrm{h}}}{U_{\mathrm{l}}}) + 0.5\log(\frac{1-U_{\mathrm{l}}h^2}{1-U_{\mathrm{l}}^2})), \tag{2.18}$$

$$A_{\mathrm{D}} = (\log\frac{C_{\mathrm{S,h}}}{C_{\mathrm{S,l}}})(3\log\frac{v_{\mathrm{h}}}{v_{\mathrm{l}}}). \tag{2.19}$$

If the DB model is a good model for eruptive solar events, the highest-velocity CME should correspond to the maximum catastrophe velocity $U_{\mathrm{h}}=1/\sqrt{2}$. This relation simplifies the second term of A_{C} yielding the phase space ratio:

$$\frac{A_{\mathrm{C}}}{A_{\mathrm{D}}} = \frac{\log(V_{\mathrm{h}}/V_{\mathrm{l}}) - 0.5\log(2-(V_{\mathrm{h}}/V_{\mathrm{l}})^2)}{3\log(v_{\mathrm{h}}/v_{\mathrm{l}})}. \tag{2.20}$$

Because only velocity ratios appear in the final expression, we replaced U (velocity in Alfven units) with $V \equiv U v_{\mathrm{A}}$ (final velocity in km s^{-1}) in all terms.

2.4.3 The CME Energy Budget

In a CME event, a large mass of plasma is lifted against gravity, is accelerated to high speed, and is generally accompanied by large bursts of radiation and showers of energetic particles. In order to account for observed CMEs, a loss of equilibrium in the DB state must provide enough energy to produce all of these events, that is,

$$E_{\rm db} = E_{\rm cme} + E_{\rm r} + E_{\rm p}. \tag{2.21}$$

The energy $E_{\rm db}$ produced by this catastrophe is $.5U^2$ in normalized units; in physical units it is $0.5M(Uv_{\rm A})^2$. The energy $E_{\rm cme}$ needed to produce the CME alone is the sum of the kinetic energy of the CME measured far from the Sun, and the gravitational potential energy difference between the solar surface and infinity

$$E_{\rm cme} = \frac{1}{2}Mv_{\rm m}^2 + \frac{G{\rm M}_\odot M}{{\rm R}_\odot}. \tag{2.22}$$

Here $v_{\rm m}$ is the measured velocity of a CME far from the sun. The extra-CME energy is in the form of radiation and may be estimated as $E_{\rm r} \approx .5E_{\rm cme}$ (Emslie et al., 2005; Dennis et al., 2006). The energy carried by high energy particles, accelerated by the CME shocks, has been estimated to be $E_{\rm p} \approx 0.1E_{\rm cme}$(Mewaldt, 2005). Using these estimates, the resulting equation for the catastrophe velocity V in terms of the observed CME velocity $v_{\rm m}$ becomes

$$V(kms^{-1}) = \sqrt{1.6(v_{\rm m}^2 + (873\ kms^{-1})^2)}. \tag{2.23}$$

This equation is really a test for the validity of the DB model to simulate the CMEs; the measured velocity $v_{\rm m}$ pertaining to observed CMEs and critical V predicted for DB states must be approximately equal.

2.4.4 Velocity Ranges for the Initial and Final States

The phase space ratio can be calculated using the minimum and maximum values of the velocity of precursors, and the minimum and maximum values of the peak CME velocities. As mentioned above, the maximum velocity of coronal flows in active regions is approximately 100 km s^{-1}. Determining a minimum velocity for coronal flows is difficult, because the uncertainty in Doppler measurements is generally on the order of 5 km s^{-1} for the strongest lines. We will conservatively assume that the minimum velocity of coronal flows is 5 km s^{-1}; thus, the range of initial speeds is 5-100 km s^{-1}. CME velocities range from about 100-1500 km s^{-1}, with a mean velocity of 487 km s^{-1} and a long high-energy tail (Schwenn et al., 2006). We omit the few CMEs in this tail and the extremely low-velocity CMEs with $v_{\mathrm{m}} < 100$km s^{-1}, attributing their range to that of the Alfven speed. Applying the CME energy budget equation, we find that the range of V is 1100-2200 km s^{-1}. There is a remarkable overlap in the observed and predicted ranges.

2.4.5 Prediction of the CME rate

The phase space ratio, measuring the fraction of catastrophe-prone states, can now be calculated and is found to be .046. The resulting predicted CME rate is 11/day, higher than the true rate of 6 d^{-1} during solar maximum (Schwenn et al., 2006). However, the systematic uncertainties may be responsible for this discrepancy. The true range of velocities in the initial state is likely to be larger than that provided by current observations, because they cannot observe flows that are too slow, or those that vary over relatively small regions. In addition, some of these DB catastrophes may result in a confined flare rather than a CME. Therefore, the predicted CME rate is in reasonable agreement with observations.

2.5 Conclusions and Discussion

In this paper, we have investigated whether solar eruptions could be produced by a catastrophic breakdown of the Double Beltrami equilibrium of Hall MHD. Our main conclusions are:

1. A catastrophe can occur in DB states that satisfy typical coronal conditions (ie, a magnetically dominated state, with large scale sizes much larger than the Hall scale).
2. This catastrophe converts up to half of the magnetic energy into kinetic energy; the released energy was shown to be large enough to drive a solar flare or CME.
3. The model predicts that the DB state's scale size expands by a factor of 2 or less as it evolves toward catastrophe, consistent with the kinematics of the quasi-equilibrium phase of LASCO C1 CMEs.
4. The predicted probability of an active region state being CME-prone is 0.046; Taking active regions to correspond to DB states, and estimating that new states appear every hour, this corresponds to a CME rate of 11 d^{-}1, which is reasonably close to the actual rate of 6 per day.

The DB model is extremely robust; for any region that is magnetically dominated, DB states will be the most likely equilibria if the Hall effect is more important than dissipation. Such regions include the magnetospheres of stars and accretion disk coronae. If observations indicate the existence of magnetically dominated, near-force-free states on a certain object, these observations are consistent with the presence of DB states. The DB model then provides a channel for eruptive or flaring events through the catastrophic breakdown of equilibrium.

Future work can make the DB model more physically realistic and detailed, allowing detailed comparison to observations. Because it is an equilibrium model, the DB model cannot predict the final destination of the energy released by the catastrophe or the detailed features of the post-equilibrium evolution. It is relatively easy to generalize the constant-density, purely magnetic model to compressible cases with an equation of state which include the effects of pressure variation and gravity (OSYM). By making this generalization and choosing an initial field geometry, it should be possible to determine if the DB catastrophe results in a CME, a confined flare, or something else in a particular physical situation; this model can then be compared directly with observed solar events.

The DB equilibrium equations could also be applied directly to other stellar coronae, as well as to accretion disk coronae, which have similar sub-Alfvenic flows and are magnetically dominated (Galeev et al., 1979). The Double Beltrami equations provide a framework for creating detailed models of eruptive events on the Sun and other astrophysical objects.

Chapter 3

A Flux Rope Network and Particle Acceleration in Three Dimensional Relativistic Magnetic Reconnection[1]

3.1 Introduction

Magnetic reconnection (e.g., Yamada et al., 2010, and references therein) is of interest in diverse areas of astrophysics, yet its mechanics remains incompletely understood. A tearing instability is thought to be necessary to initiate reconnection in reversing magnetic field configurations, but many astrophysical plasmas have collisional resistivities that are insufficient to facilitate its growth. Interpretations of space plasma measurements (e.g., Chen et al., 2008; Øieroset et al., 2011) and astronomical observations, however, suggest that efficient collisionless reconnection is ubiquitous. Therefore, collisionless effects, which operate on plasma kinetic scales, seem to be required to provide the dissipation necessary for effecting a change of magnetic topology. Magnetohydrodynamic (MHD) models for the dynamics of systems undergoing magnetic reconnection have been available for a long time (e.g., Priest & Forbes, 2000, and references therein). Unfortunately, these models do not describe the underlying nature of (possibly multiscale) plasma organization in the

---

[1] A large part of the material in this chapter is taken from the manuscript Kagan, D., Milosavljević, M., and Spitkovsky, A., 2013, which has been accepted for publication by the Astrophysical Journal.

reconnection layer where magnetic energy is being dissipated and the assumptions of ideal MHD do not apply. Understanding the detailed plasma organization on all length scales, from the likely relatively small, plasma kinetic scales, to the potentially much larger scales on which astrophysical dynamical systems “prepare” reconnection sites, and where ideal MHD may be valid, is paramount for completing the theories of a wide variety of astrophysical phenomena and for interpreting space plasma measurements and astronomical observations.

In an effort to develop a picture of magnetic reconnection from first principles, recent particle-in-cell (PIC) simulations have examined the dynamics of reconnection layers that start with a current sheet as thin as the plasma skin depth. PIC simulations of reconnection in pair plasmas have been carried out in two spatial dimensions (Zenitani & Hoshino, 2001, 2007; Jaroschek et al., 2004; Bessho & Bhattacharjee, 2005, 2007, 2010, 2012; Daughton & Karimabadi, 2007; Hesse & Zenitani, 2007; Hoshino, 2012) and three dimensions (Zenitani & Hoshino, 2005, 2008; Yin et al., 2008; Liu et al., 2011; Sironi & Spitkovsky, 2011). Among these, several (Zenitani & Hoshino, 2005, 2008; Bessho & Bhattacharjee, 2007; Hesse & Zenitani, 2007; Hoshino, 2012) have investigated the role of departure from the idealized, exactly antiparallel reconnection by introducing a perpendicular “guide” field. These various simulations have revealed novel forms of small-scale plasma self-organization that are interesting in their own right, but that must ultimately be related to and embedded within the appropriate larger astrophysical contexts (e.g., Uzdensky et al., 2010). While spacecraft measurements, which can be done in situ, can provide direct clues how to establish this embedding in space plasmas, in extrasolar contexts only an indirect relation can be established between the reconnection process and the observed emission (e.g., Sironi & Spitkovsky, 2011; Cerutti et al., 2012b).

Common features seen in many PIC simulations of magnetic reconnection include the formation of chains of magnetic flux ropes (in three dimensions with a guide field; otherwise, the common terms "islands" or "plasmoids" may still be more appropriate), the merging of smaller flux ropes into larger ones, and an energization of the plasma in the reconnection layer. In three dimensional simulations, kink-like and oblique modes, as well as secondary instabilities, can impart three dimensional structure to the reconnection layer.

Typically, the simulations are initialized in the so-called Harris sheet equilibrium describing a current sheet with a thickness similar to the plasma skin depth. The tearing instability first sets in on scales of the initial current sheet thickness. Its nonlinear development produces a chain of skin-depth-scale flux ropes alternating with magnetic X-lines, the three dimensional generalization of two dimensional X-points. In X-lines, violation of flux freezing and magnetic line reconnection can be facilitated by a pressure tensor anisotropy (Vasyliunas, 1975, see also, e.g., Hesse & Zenitani 2007, and references therein). Smaller flux ropes tend to merge with each other to form larger ones; this gives rise to magnetic organization on increasingly larger spatial scales. Three-dimensional PIC simulations of guide field reconnection in electron-ion plasmas exhibit these same features, e.g., Daughton et al. (2011) found that oblique modes dominated over tearing modes when guide field was strong. Additional effects specific to plasmas with electron-ion mass disparity have also been identified, but are not relevant for the present work.

Energization of particles in reconnection layers has been investigated in a number of PIC simulations (Zenitani & Hoshino, 2001, 2007; Jaroschek et al., 2004; Drake et al., 2006, 2010; Bessho & Bhattacharjee, 2007, 2010, 2012; Egedal et al., 2009; Huang et al., 2010; Oka et al., 2010b; Liu et al., 2011; Sironi & Spitkovsky, 2011; Egedal et al., 2012; Hoshino, 2012; Cerutti et al., 2012b). Less attention has been given to particle energization in the general case of reconnection with a guide field in three dimensions (Zenitani & Hoshino, 2008). In two dimensional simulations, particle acceleration producing a nonthermal energy spectrum, an apparent power-law, is often reported. Cerutti et al. (2012b), however, instead detect a new ultrarelativistic thermal component energized by the reconnection.

That the reconnection layers should energize particles is in agreement with analytical considerations (e.g., Speiser, 1965; Larrabee et al., 2003; Giannios, 2010; Uzdensky et al., 2011; Cerutti et al., 2012a), which find that particles in the vicinity of the X-line in the reconnection layer are accelerated by the nearly-uniform electric field as they repeatedly cross, and are trapped within the converging plasma flows. Other mechanisms focusing on energetic particles that have moved from the X-line region into the flanking islands have also been suggested (e.g., Drake et al., 2006, 2010). In three dimensional simulations, evidence for a nonthermal spectrum is less solid. It remains poorly understood which processes limit the energy to which particles can be accelerated in fully dynamical, three-dimensional reconnection layers, and how do the particle energy spectrum, the degree of accelerated particle beaming, and the temporal evolution of the accelerated population depend on the parameters of the reconnection layer.

In this work we employ three dimensional PIC simulations to investigate the evolution of current sheets in relativistic pair plasmas undergoing magnetic reconnection. Our simulations add to the small but growing family of three dimensional PIC simulation of relativistic reconnection with a guide field. With the intention to complement existing work, we initialize our simulations slightly differently than it is normally done, not assuming the usual Harris sheet equilibrium. Also, we explore a parameter regime, involving magnetic to kinetic pressure ratios of the order of unity, that has thus far not received sufficient attention. We observe an evolution of magnetic field geometry that constrains the viability of models for high-Lundquist-number reconnection layers in which the diffusion region contains a hierarchy of interacting plasmoids (e.g., Shibata & Tanuma, 2001; Uzdensky et al., 2010). The simulations also allow us to explore the character of particle energization in dynamical, fully three dimensional reconnection layers.

This paper is organizes as follows. Section 3.2 describes our methodology and simulation setup, while Section 4.3 presents the results. Section 3.4 discusses our findings concerning development of kinetic instabilities in the current sheet, as well as our findings on particle energization, in view of the existing work on these topics. Finally, Section 3.5 reviews our main conclusions.

3.2 Description of Simulations

3.2.1 The Initial Configuration

The spatial domain is a rectangular box with $0 \leq x < L_x$, $0 \leq y < L_y$, and $0 \leq z < L_z$. The boundary conditions are periodic in all directions. The initial magnetic field is the same as that in the Harris equilibrium

$$\begin{aligned}\mathbf{B} &= B_0\left[\tanh\left(\frac{x-L_x/4}{\lambda_0}\right)-\tanh\left(\frac{x-3L_x/4}{\lambda_0}\right)-1\right]\hat{\mathbf{z}}\\ &\quad +\kappa B_0(-\hat{\mathbf{y}}),\end{aligned}\tag{3.1}$$

where λ_0 is the half-width of the initial current sheet and $\kappa \geq 0$ is a parameter defining the strength of the uniform guide field perpendicular to the opposing field, which in our simulations is oriented in the $-y$ direction. The current sheets are located at $x = L_x/4$ and $x = 3L_x/4$ and carry antiparallel currents. The current density profile that satisfies Ampère's law is

$$\mathbf{J} = -\frac{cB_0}{4\pi\lambda_0}\left[\operatorname{sech}^2\left(\frac{x-L_x/4}{\lambda_0}\right)-\operatorname{sech}^2\left(\frac{x-3L_x/4}{\lambda_0}\right)\right]\hat{\mathbf{y}}. \tag{3.2}$$

To ensure that Ampère's law is satisfied, if we set the particle density to be uniform, the particles have a spatially-dependent drift velocity $\boldsymbol{\beta}_i = -\boldsymbol{\beta}_e = \boldsymbol{\beta}$ in the $-y$-direction. The current density $\mathbf{J}$ is related to the velocity $\boldsymbol{\beta}$ by the relation

$$\mathbf{J} = n_0 ec(\boldsymbol{\beta}_i - \boldsymbol{\beta}_e) = 2n_0 ec\boldsymbol{\beta}, \tag{3.3}$$

where e is magnitude of the unit charge carried by electron and ion macroparticles. This results in the drift velocity profile

$$\boldsymbol{\beta} = -\beta_0\left[\operatorname{sech}^2\left(\frac{x-L_x/4}{\lambda_0}\right)-\operatorname{sech}^2\left(\frac{x-3L_x/4}{\lambda_0}\right)\right]\hat{\mathbf{y}}, \tag{3.4}$$

where $\beta_0 \equiv B_0/(8\pi n_0 e\lambda_0)$.

In this work, we initialize the simulation outside of pressure equilibrium, which is a significant difference from other works; we begin with a uniform initial density n_0 and spatially varying drift velocity $\boldsymbol{\beta}$. We can still attempt to relate the parameters of our simulation to those of preceding investigations. In setting up initial conditions for a plasma with the magnetic field given in Equation (3.1), there are multiple ways in which pressure equilibrium can be satisfied depending on the spatial variation of the plasma density $n_{e^+}+n_{e^-}$, temperature T, charge drift velocity $\boldsymbol{\beta}$, reversing field strength B_0, and guide field strength B_y. It is common to assume that the initial temperature is uniform and equal to T_0, and only the density varies across the current sheet. In practice, Harris sheets are often set up with a strong excess density in the current sheet for pressure balance. It is common to split the particle population into two components, one a uniform background with density n_{b}, and another spatially varying with maximum density n_0 at the center of the current sheet. The latter component ensures pressure equilibrium and carries the current in the reconnection layer. In Section 3.2.4 below, we discuss how the initial configuration adjusts to approximate pressure equilibrium.

One can define the magnetic to kinetic pressure ratio via

$$\sigma \equiv \frac{P_{\mathrm{mag}}}{P_{\mathrm{kin}}} = \frac{B^2}{8\pi(n_{e^+}+n_{e^-})T}, \tag{3.5}$$

where here and henceforth we express the temperature in energy units. Then in the isothermal Harris sheet pressure equilibria used in preceding investigations, the ratio simply equals the density contrast in the current sheet, $\sigma = n_0/n_{\mathrm{b}}$. In our simulation, σ is uniform outside the current sheet and equals

$$\sigma = \frac{B_0^2}{16\pi n_0 T_0}. \tag{3.6}$$

Note that σ is defined not taking into account the magnetic pressure of the guide field.

We do not introduce any initial perturbation to the initial field geometry described here. The structure that develops is thus seeded by numerical fluctuations in conjunction with the initial pressure imbalance.

3.2.2 Parameters

We initialize the simulation at temperature $T_0 = m_e c^2$. All the particles are drawn from the relativistic Maxwellian distribution, implying that the average kinetic energy of the particles is $\sim 2.37\, m_e c^2$. We are interested in the dependence of reconnection mechanics and evolution of particle energy distribution on the dimensionless ratio of the magnetic pressure to the particle pressure σ, and the guide field amplitude parameter κ. A nonvanishing κ indicates that the magnetic field is twisted in the current sheet and the plasma in the sheet center is magnetized. We run a grid of nine simulations, with $0.25 \leq \sigma \leq 2$ and $0 \leq \kappa \leq 1$, and carry out a detailed study of the run with $\sigma = 2$ and $\kappa = 0.25$. The chosen values of σ are low because we are interested in finding the lower limit of magnetization at which reconnection can produce significant particle energization.

Because of the low growth rate, the simulations with $\sigma = 0.25$ did not develop the tearing instability over the time period of the simulations and thus did not undergo reconnection, therefore, we do not show the results of these simulations in what follows. The parameters of the simulations that did undergo reconnection are shown in Table 3.1.

Table 3.1. Simulation Parameters and Results

Run	L_x [a]	L_y [a]	L_z [a]	λ_0 [a]	σ	κ	f_c [b]	$\lvert\Delta\mathcal{E}_B\rvert/\mathcal{E}_B$ (%) [c]	$K_{\rm ener}/K$ (%) [d]	$K_{\rm ener}/\lvert\Delta\mathcal{E}_B\rvert$ (%) [e]	$\gamma_{\rm max}$ [f]
`S1K0`	64	40	40	2	1	0	1.5	72	3.1	11	48.4
`S1K025`	64	40	40	2	1	0.25	1.0	28	3.1	23	34.3
`S1K1`	64	40	40	2	1	1	1.0	4.3	1.5	32	34.2
`S2K0`	64	60	60	3	2	0	2.3	55	10	30	52.6
`S2K025`	64	60	60	3	2	0.25	2.3	44	9.3	29	54.4
`S2K025L`	128	120	120	3	2	0.25	2.3	27	9.1	38	58.9
`S2K05`	64	60	60	3	2	0.5	2.3	25	6.4	27	44.9
`S2K1`	64	60	60	3	2	1	2.3	6.7	3.2	28	42.8

[a] The length scales L_x, L_y, L_z, and λ_0 are given in units of the plasma skin depth $\lambda_{\rm p}$.

[b] $f_{\rm c}$ is the ratio of the original current sheet width to the current sheet width after the readjustment phase.

[c] The ratio $|\Delta\mathcal{E}_B|/\mathcal{E}_B$ is the fraction of magnetic energy converted into kinetic energy during the simulation.

[d] $K_{\rm ener}/K$ is the particle kinetic energy fraction in energized particles (see Section 3.3.7.1).

[e] $K_{\rm ener}/|\Delta\mathcal{E}_B|$ is the ratio of the particle kinetic energy in energized particles to converted magnetic energy (see Section 3.3.7.1).

[f] $\gamma_{\rm max}$ is the Lorentz factor of the highest energy particle in the simulation.

3.2.3 Simulation Method and Resolution Requirements

We use the relativistic PIC plasma code `TRISTAN-MP` (Spitkovsky, 2008) to simulate the evolution of a reconnection configuration in a pair plasma. In a PIC simulation, the number of macroparticles of each species located in each grid cell must be large enough to resolve variations in the current density and limit high-frequency particle noise. `TRISTAN-MP` uses a current filtering algorithm to reduce high-frequency particle noise, substantially reducing the required number of macroparticles per cell per species. We initialize our simulations with 4 macroparticles per cell per species, which is fewer than found in other recent PIC simulations (Daughton et al., 2011; Liu et al., 2011). To verify that this low particle density does not lead to cell evacuation, we calculate the total particle density including both species in each cell in Run `S2K025` during the flux rope merging phase discussed in Section 3.3 below; this is the run and time at which the smallest densities are present. We find that the total particle density in each cell calculated with a cubic cloud-in-cell kernel of size equal to the grid spacing is almost always larger than 1, with a total density larger than 5 in over 93% of cells. Cells with low particle density occur in regions characterized by weak spatial field gradients that should not be sensitive to particle noise. The cells with strong field gradients that would be sensitive to particle noise invariably have a higher than average particle density.

To ensure that the particle density is indeed sufficient to resolve the physics of reconnection, we have carried out two dimensional convergence tests, as well as a longer three dimensional test run with 80 particles per cell per species. In our simulations, the tearing instability leading to reconnection is seeded by numerical fluctuations that vary from simulation to simulation, so we do not expect all observables to be the same in each run at a given absolute time measured from the beginning of the simulation. We therefore compare measurable quantities in each simulation at similar points in their evolution; typically, we calculate these quantities at the end of the flux rope merging phase, which is discussed in Section 3.3 below. The coarse time discretization employed in recording simulation output makes the determination of these evolution points imprecise. This results in inaccuracies in the estimation of observables that vary quickly with time; such quantities will therefore vary with particle density even if the resolution in particle density is sufficient to capture the physics of reconnection.

Our two dimensional convergence tests use initial conditions matching those in the three dimensional simulations with $\sigma = 2$ and $\kappa = 0$, but with a larger number, up to 16, of macroparticles per species per cell. We find that the only significant discrepancy, up to 40%, is seen in the peak linear growth rate of the kinetic tearing instability; this quantity is very sensitive to the coarse time discretization because time differencing is required for growth rate calculation. This discrepancy is reduced to 25% in simulations with somewhat finer time discretization in recording outputs; therefore, insufficient particle density is not responsible for the discrepancy in tearing mode growth rates. To test that we have sufficient macroparticle density to account for 3D effects in our simulations, we also carry out a three dimensional simulation with 80 macroparticles per cell per species, and the same parameters as in Run `S2K025` with $\sigma = 2$ and $\kappa = 0.25$. We again find that the measurable quantities do not deviate greatly from those found in Run `S2K025`, with the maximum deviation of

30% again seen in the growth rate of the kinetic tearing instability. We also find that the evolution of the total magnetic energy and the particle energy spectrum in this simulation with higher macroparticle density does not differ greatly from that found in Run `S2K025` in Section 3.3.7. The main difference is that in the high-density simulation, the flux rope merging phase identified in Section 3.3 begins slightly later; this may be understood as the result of slightly lower particle noise in the high density simulations, which results in later growth of instability.

To ensure that our simulations resolve the plasma skin depth, we set $\Delta x = \lambda_\mathrm{p}/8$, where the skin depth is given by

$$\lambda_\mathrm{p} = \sqrt{\frac{\langle\gamma\rangle\, m_e c^2}{8\pi n_0 e^2}}. \tag{3.7}$$

Here, $\langle\gamma\rangle$ is the average Lorentz factor of particles in the simulation and the additional factor of 2 in the denominator reflects the fact that the electrons and positrons oscillate together. To determine whether this value for λ_p is sufficient to resolve the physics of reconnection, we have carried out further two dimensional convergence test simulations with up to 3 times larger number of grid cells per skin depth than in the three dimensional simulations, and again compared results at the end of the flux rope merging phase. Here, the most resolution–dependent quantity was the maximum value of the reconnected magnetic field B_x, which varied by 25% with resolution. This is a quantity that varies quickly with time near the end of the flux rope merging phase, so the apparent variation with resolution is again likely a result of coarse time discretization.

We focus on the regime in which the particle drift velocity in Equation (3.4) giving rise to the current density in Equation (3.2) is nonrelativistic. This places a constraint on the initial current sheet width λ_0. Combining Equations (3.2) and (3.3) with the definition of σ in Equation (3.6), it can be shown that

$$\frac{\lambda_0}{\lambda_\mathrm{p}} = \frac{\sqrt{2\sigma(\Gamma-1)}}{\beta_0}, \tag{3.8}$$

where $\Gamma - 1 \equiv T_0/(\langle\gamma\rangle\, m_e c^2)$ is the ratio of the particle pressure to the particle energy density (which includes the particle rest energy); in our simulations, $\Gamma - 1 \approx 0.297$.[2] With this estimate of the relation between λ_0 and β_0, in simulations with $\sigma \leq 1$ we choose $\lambda_0/\lambda_\mathrm{p} = 2$, while in simulations with $\sigma = 2$ we choose $\lambda_0/\lambda_\mathrm{p} = 3$. This implies that $\beta_0 \approx 0.36$, ensuring that the drift velocities are not relativistic.

3.2.4 Readjustment to Equilibrium

The uniform initial density and temperature and nonuniform magnetic field in the initial configuration defined by Equations (3.1), (3.3), and (3.4) are not in force balance, but they achieve approximate force balance following a brief readjustment. The initial adjustment to equilibrium results in a reduction of the current sheet thickness and an accompanying compression of the plasma by a factor

$$f_\mathrm{c} \equiv \frac{\lambda_0}{\lambda} > 1, \tag{3.10}$$

---

[2] The index Γ is related to the temperature via

$$\frac{1}{\Gamma-1} = \frac{1}{\theta}\left[\frac{K_3(\theta^{-1})}{K_2(\theta^{-1})} - \theta\right], \tag{3.9}$$

where $\theta \equiv T/(m_e c^2)$ and K is the modified Bessel function of the second kind.

where henceforth, λ denotes the current sheet half-thickness after the readjustment. We further define

$$\tau_{c0} \equiv \frac{\lambda_0}{c}, \quad \tau_c \equiv \frac{\lambda}{c} \tag{3.11}$$

to denote the light crossing times of the initial and the compressed current sheet.

The plasma takes $\approx 10\tau_{c0}$ to adjust to a state close to pressure equilibrium. The magnitude of the compression is shown in Table 3.1. With readjustment, both the density of the particles and the strength of the guide field increase in the center of the current sheet in such a way as to balance the large pressure gradient that the reversing field exhibits from the center to the surface of the current sheet. Longer-term transients are present in the form of acoustic oscillations; we show that these are magnetosonic waves in Section 3.3.1. These oscillations could have the effect of enhancing the tearing instability, which relies on a velocity inflow to the current sheet. The initial readjustment does not give rise to any significant magnetic dissipation or particle energization.

3.2.5 Unstable Modes and the Size of the Simulation

For a simulation to capture the physics of reconnection, its size must be sufficient to include the fastest growing modes of the important instabilities of relativistic Harris current sheets in pair plasmas. The known instabilities include the kinetic relativistic tearing instability (KTI; hereafter the tearing mode), the drift-kink instability (DKI; hereafter the kink mode), and the oblique mode which is similar to the tearing mode. To estimate their wavelengths, we follow Zenitani & Hoshino (2007). The growth rate ω_{KTI} of the tearing mode with wave number k_z in a Harris current sheet of half-thickness λ and non-relativistic drift velocity β is given by

$$\omega_{\mathrm{KTI}} = b(T)\,\beta^{3/2}\,k_z\lambda\,[1-(k_z\lambda)^2]\,\frac{1}{\tau_c}, \tag{3.12}$$

where is a dimensionless function of the plasma temperature that in the limits of a cold and a relativistically hot plasma equals

$$b(T) = \begin{cases} \frac{1}{\sqrt{\pi}}\left(\frac{2T}{\gamma_\beta m_e c^2}\right)^{-1/2} & ,\ T \ll m_e c^2, \\ \frac{2\sqrt{2}}{\pi} & ,\ T \gg m_e c^2, \end{cases} \tag{3.13}$$

and $\gamma_\beta \equiv (1-\beta^2)^{-1/2}$. The resulting maximum growth rate occurs for $k_z\lambda = 1/\sqrt{3}$, corresponding to a wavelength of $\sim 10.8\,\lambda$. If we use the form of $b(T)$ appropriate at ultrarelativistic temperatures, the growth rate for this mode is

$$\omega_{\mathrm{KTI,max}} = 0.35\,\frac{\beta^{3/2}}{\tau_c}. \tag{3.14}$$

The situation is more complicated for the kink mode, because as Zenitani & Hoshino (2007) find, the analytical maximum value of the growth rate ω_{DKI} occurs for $k_y\lambda > 1$, a wavenumber for which kinetic effects become important in a thin current sheet. Simulation results in their Figure 20 indicate that the maximum growth rate occurs at $k_y\lambda \approx 0.7$ corresponding to a wavelength of $\approx 9\,\lambda$.

In addition to these two dimensional modes, it is also necessary to resolve oblique modes with $k_y, k_z \neq 0$ that combine tearing and kink components; these were identified in three dimensional simulations by Zenitani & Hoshino (2005, 2008) and Daughton et al. (2011). The typical fastest-growing oblique mode in both of these simulations had $k\lambda \sim 0.2$; this corresponds to a wavelength in both directions of $\sim 30\,\lambda$.

To resolve all three types of modes, in all simulations except for the larger size run `S2K025L`, we set $L_y = L_z = 20\,\lambda_0$, large enough to contain at least one wavelength of the tearing mode and two wavelengths of the kink mode. Because the initial adjustment leads to a significant narrowing of the current sheet, several wavelengths of the fastest-growing tearing and drift-kink modes of the narrower current sheet are included in the simulations, and at least one wavelength of the oblique modes should also be resolved. The resulting overall length scales are $L_x = 64\,\lambda_\mathrm{p}$ and $L_y = L_z = 40\,\lambda_\mathrm{p}$ or $L_y = L_z = 60\,\lambda_\mathrm{p}$, depending on the simulation, as is shown in Table 3.1. To ensure that little interaction takes place between the two current sheets in the periodic box, we set $L_x = 64\,\lambda_\mathrm{p}$. The current sheets in our larger size simulation `S2K025L` are as thick as in the other ones, but their separation and the dimensions of the box are twice as large, i.e., $L_x = 128\,\lambda_\mathrm{p}$ and $L_y = L_z = 120\,\lambda_\mathrm{p}$. This simulation was carried out on a $1024 \times 960 \times 960$ grid and contained a total of $\approx 7.5 \times 10^9$ particles. The smaller simulations with $\sigma = 1$ were carried out on a $512 \times 320 \times 320$ grid and contained 4.2×10^8 particles, whereas the simulations with $\sigma = 2$ were carried out on a $512 \times 480 \times 480$ grid with 9.4×10^8 particles.

3.2.6 Duration of the Simulations

In our simulations the growth times of the tearing and kink modes are typically similar to or smaller than the Alfvén crossing time of the box. To capture the physics of reconnection in its nonlinear regime and on length scales similar to the box size, the duration of the simulation must be larger than these time scales.

The relativistic Alfvén velocity is given by

$$v_{\mathrm{A}} = \frac{c}{\sqrt{1+\Gamma/[2\sigma(\Gamma-1)]}}. \tag{3.15}$$

The Alfvén crossing time of the box $\tau_{\mathrm{A},z} = L_z/v_{\mathrm{A}}$ is $\sim 28\,\tau_{c0}$ in the runs with $L_z = 60\,\lambda_{\mathrm{p}}$. We run all of our simulations for at least 8000 time steps, which amounts to $150\,\tau_{c0}$ in the large simulation `S2K025L`.

3.3 Results

The evolution of the reconnection layer in all simulations exhibits the same common properties that have been observed in previous PIC simulations of magnetic reconnection in pair plasmas. The tearing instability grows and produces a chain of alternating X-lines and flux ropes. Here, we adopt “flux ropes” to denote a magnetic structure sometimes referred to as “plasmoid” (or “island” in two dimensional treatments) in which plasma is pinched by a helical magnetic field which can have a braided structure. The flux ropes can have a finite length, with the field lines opening up and extending arbitrarily far from the rope axis. The flux ropes can also split into sub-ropes, which opens the possibility of an organization of the ropes in a flux rope network. In all simulations except for `S1K0` where the kink instability disrupts the current sheet, the flux ropes merge in quasi-hierarchical fashion until either only one flux rope is left in each current sheet or, in the large simulation `S2K025L`, the flux ropes have been disrupted by a transition to a disordered, three-dimensional state. The period of flux rope merging is accompanied by fast magnetic-to-kinetic energy conversion. This conversion produces a tail of energized particles.

In simulations without guide field, the kink instability also grows, resulting in some corrugation of the current sheet. For $\sigma = 2$ in simulation `S2K0`, the kink instability does not disrupt the quasi-hierarchical merging process. For $\sigma = 1$ in simulation `S1K0`, however, the kinking disrupts the merging and brings the two current sheets into contact where they can interact; this is accompanied by rapid conversion of magnetic to thermal energy but the resulting spectrum does exhibit signatures of a secondary energized component. In agreement with Zenitani & Hoshino (2008), we find that the presence of a guide field suppresses the kink instability. However we find that in the large simulation `S2K025L`, a transition to three-dimensional evolution still occurs, likely due to a lack of large-scale phase coherence in the tearing instability.

We proceed to discuss our results in detail. In Section 3.3.1 we discuss the evolution of the global current sheet and the formation of what will turn out to be a network of interconnected flux ropes in our largest simulation. In Section 3.3.2 we identify sites of magnetic reconnection within the network. In Section 3.3.3 we carry out a Fourier decomposition of perturbations in the largest simulation and discuss their growth rates. In Section 3.3.4 we analyze the time scales associated with the evolution of the flux rope network. In Section 3.3.5 we measure the overall reconnection rate, while in Section 3.3.6 we analyze components of the nonideal electric field as reflected in the generalized Ohm's law. Finally, in Section 3.3.7 we discuss the overall rate of magnetic-to-thermal energy conversion as well as the efficiencies, mechanisms, and properties of particle energization in the simulations.

3.3.1 Formation and Evolution of the Flux Rope Network

Figure 3.1 shows a time sequence of slices at $y = 5\,\lambda_{\rm p}$ of the total current density $J \equiv |\mathbf{J}|$ and the plasma number density n in simulation `S2K025L`. The early evolution of each current sheet, if observed at a single value of y, is that correctly described by the familiar flux rope (or island, plasmoid) merging paradigm. Larger flux ropes produced by the merging of smaller ones contain substructure reflected in multiple, curved, embedded current sheets, each with a half-thickness $\sim \lambda_{\rm p}$. Very similar substructure also appears in the recent two-dimensional PIC simulations of an electron-ion plasma by Markidis et al. (2012).

Outside of the two evolving current sheets, plane-parallel magnetosonic waves resulting from the initial pressure imbalance can be seen in the plasma density, the electric field in the y direction, and the magnetic field in the z direction. The waves traverse the computational box several times in the course of the simulation, moving with a typical speed of approximately $0.8c$. To check that these are indeed magnetosonic waves, we calculate the magnetosonic speed $v_{\rm ms} = c\sqrt{(v_{\rm A}^2 + c_{\rm s}^2)/(v_{\rm A}^2 + c^2)}$, where $c_{\rm s}$ is the local sound speed given by $c_{\rm s} = c\sqrt{\Gamma - 1} \sim 0.55c$. Because the typical density in the background plasma is $\sim 0.8n_0$ and a guide field of magnitude one quarter that of the reversing field is present, the Alfvén speed in the background plasma is approximately $0.73c$. The magnetosonic speed in the background plasma is therefore approximately $0.74c$, which is similar to the speed of the observed waves.

To examine three dimensional aspects of the current sheet evolution, in Figure 3.2 we show projections into the yz plane of the total current density J,

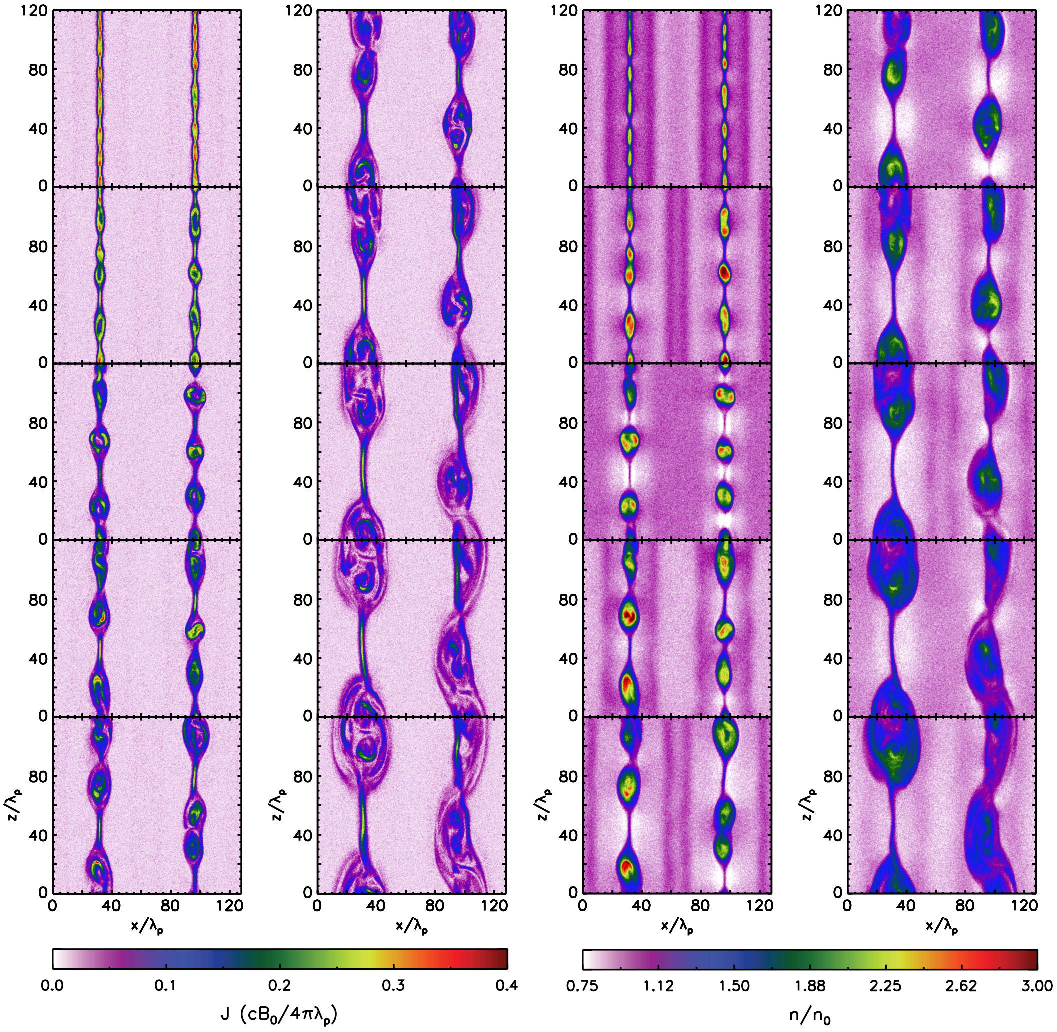

Figure 3.1 The total current density $J \equiv |\mathbf{J}|$ (left two columns) and total particle density n (right two columns) in a slice $y = 5\,\lambda_\mathrm{p}$. In each pair of columns, the first column shows times $t = (37.5, 47, 56, 65.5, 75)\,\tau_{c0}$ and the second column shows $t = (84.5, 94, 103, 112.5, 122)\,\tau_{c0}$. The particle density n, which is not used directly in the PIC code, is calculated from particle positions using a uniform kernel of half-width 2 grid lengths. The same kernel is used to smooth the current density J.

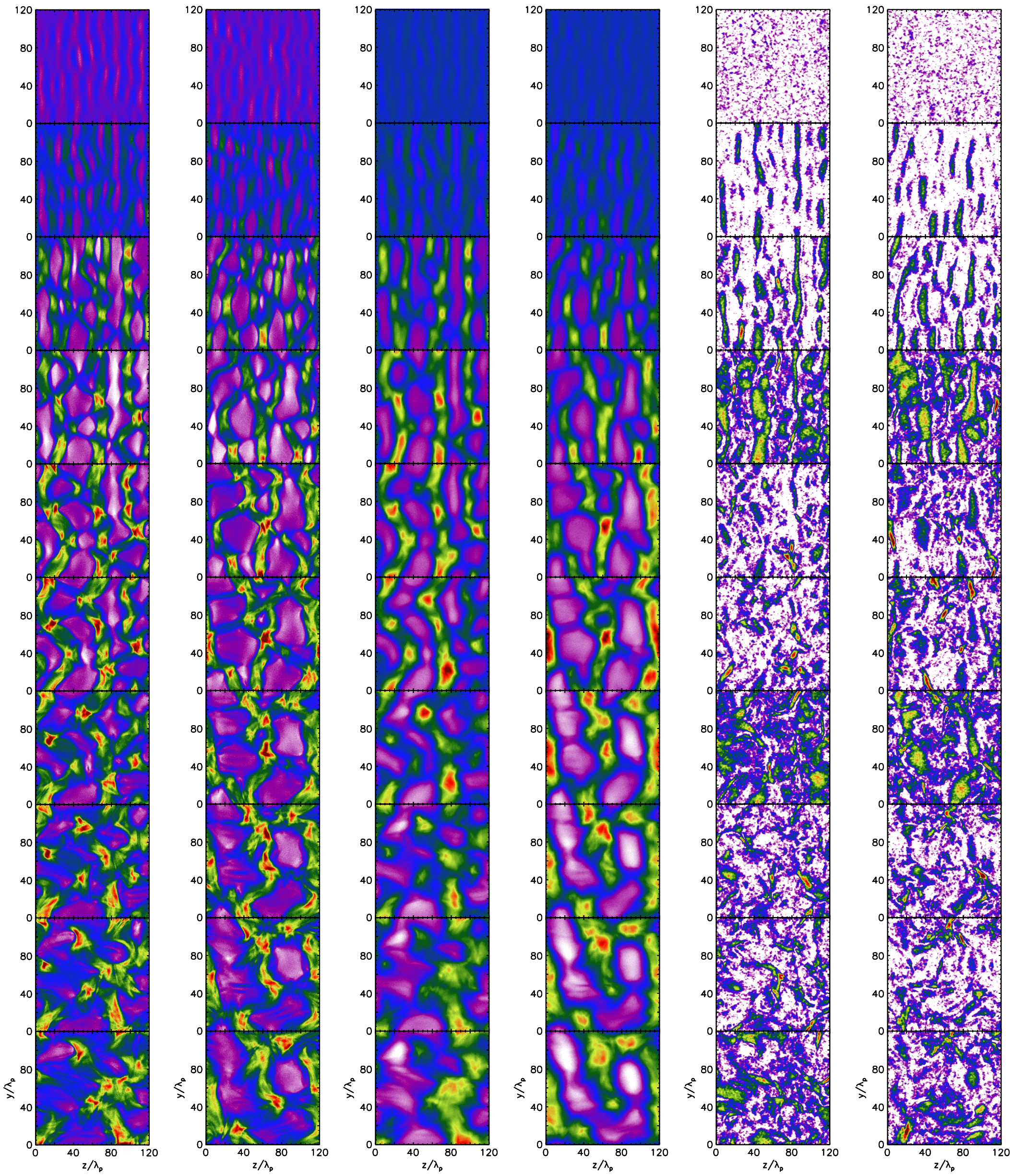

Figure 3.2 The total current density $j \equiv |\mathbf{j}|$ (left two columns), total particle density n (middle two columns), and the parallel electric field magnitude $E_{||} \equiv |\mathbf{E}\cdot\mathbf{B}|/B$ projected onto the yz plane. In each pair of columns, the left column shows a projection of the sections $0 < x < 64\,\lambda_{\mathrm{p}}$ containing one current sheet, and the right shows a projection of the section $64\,\lambda_{\mathrm{p}} < x < 128\,\lambda_{\mathrm{p}}$ containing the other current sheet. Vertically from top to bottom, the panels show times $t = (28, 37.5, 47, 56, 65.5, 75, 84.5, 94, 103, 112.5)\,\tau_{c0}$. The color scale increasing from light purple to dark red is linear except for the parallel electric field, where the scale is logarithmic. The color scale coverage of each projected quantity was reduced from the full variation of that quantity for enhanced visual contrast. In particular, clipping of the color scale for the parallel electric field excludes the range in which the parallel electric field projection is dominated by small-scale electrostatic fluctuations in the plasma. We boxcar smoothed $\mathbf{E}$ and $\mathbf{B}$ on scales $\lesssim\lambda_{\mathrm{p}}$ prior to

plasma number density n, and the parallel electric field magnitude $E_{\|}$ in each of the two current sheets in simulation `S2K025L`. The flux ropes that initially develop from the linear tearing instability are approximately parallel to the y axis, which is the direction of the initial current flow. However, each flux rope exhibits one or more discontinuities where the z coordinate of the rope suddenly changes. We interpret these discontinuities as arising from a lack of tearing mode phase coherence on scales larger than those in causal contact during the initial development of the instability. The lack of phase coherence should occur on scales

$$\Delta y, \Delta z \gg \lambda_{\mathrm{coh}} \sim \frac{c}{\omega} \tag{3.16}$$

on which locations on the current sheet are out of causal contact during the first e-folding of the instability; here, ω is the linear growth rate of a mode growing from numerical noise. This sets a hard upper limit on the scale on which an instability can grow coherently; other constraints may limit this scale further.

The initial departure from perfect translational invariance in the y direction implies that the nonlinear flux rope merging will itself not occur coherently. In any three adjacent flux ropes, the middle rope can merge with one of the flanking ropes at one y, and with the other rope at another y; at still other values of y, the three ropes may remain separate for a time being. This creates lateral linkage between the flux ropes; already at time $t = 37.5\,\tau_{c0}$, all the flux ropes are mutually linked inside their current sheet. The linkage implies each flux rope experiences a strongly y-dependent magnetic tension force. This magnetic tension from neighboring flux ropes leads to rope tilting and kinking that is very distinct in origin from that arising from the well-known linear oblique and kink instabilities of a current sheet.

Pairs of large flux ropes emerging from two generations of flux rope merging, but still oriented largely parallel to the y axis, are often connected by minor flux ropes that are highly tilted toward the z axis. Daughton et al. (2011) have discovered similar structures in a large three-dimensional simulation of electron-ion reconnection with a strong guide field. It is clear in Figure 3.2 that by $t \sim 75\,\tau_{c0}$, all flux rope segments have substantial tilts and any semblance of translational invariance in the direction of the original current flow is lost. Vertices of the flux rope network contain highly localized, intense current sheets and filaments. By $t \sim 122\,\tau_{c0}$, the flux ropes are largely completely disrupted and the current sheet contains a disordered network of knot- and sheet-like structures.

The force $\mathbf{J} \times \mathbf{B}$ peaks at the primary X-line current sheets as well as flux rope perimeters, but is on average much smaller in flux rope interiors. This suggests that the flux rope interiors are organized in a state of force free quasi-equilibrium constrained by a nonvanishing magnetic helicity (i.e., twisting, braiding of the field lines).

3.3.2 Dissipation and Reconnection in the Flux Rope Network

To gain further insight into the structure of the flux rope network, we examine the spatial variation of $\mathbf{E}_{||}$, the component of the electric field parallel to the magnetic field. The parallel field vanishes for the purely inductive electric field of ideal MHD and thus it can be obtained purely from the non-ideal, reconnection electric field $\mathbf{R}$ defined as the difference between the actual and the induction electric field,

$$\mathbf{R} \equiv \mathbf{E} + \frac{1}{c}(\mathbf{v} \times \mathbf{B}). \tag{3.17}$$

In high magnetic Reynolds number plasmas, the parallel field is present only in magnetic diffusion regions, and in particular, in locations where a change of magnetic connectivity is under way (e.g., Schindler et al., 1988).[3] The parallel electric field may also be instrumental in maintaining the pressure anisotropy, that is responsible for the breaking of flux freezing in the reconnection layer (see, e.g., Hesse et al., 2011).

The two rightmost columns of Figure 3.2 show that the magnitude of the parallel electric field $E_{||} \equiv |\mathbf{E}_{||}|$ strongly peaks in the thin, long current sheets containing primary X-lines and connecting magnetic flux ropes. However prior to flux rope disruption and the transition to a disordered state, the parallel field is normally very small inside magnetic flux ropes even when they contain embedded, substructure-level current sheets. An exception are flux rope segments directly undergoing dynamical transformation, such as merging or stretching, where very isolated, spatially and temporally intermittent sites of significant $E_{||}$ occasionally appear. The flux rope interiors are not undergoing pervasive, steady reconnection, in spite of their complex magnetic substructure. The intermittent $E_{||}$ is also common after the flux rope network has become disordered, indicating that the disordered network remains active by developing transient, localized dissipation events in the braided network of disrupted flux ropes that effect further evolution of magnetic connectivity.

Reconnection intermittency is also evident in Figure 3.3, showing views

---

[3] On very small scales in the simulation, $\mathbf{E}_{||}$ exhibits noise in the entire domain due to small scale electrostatic fluctuations; in our analysis, we filter these fluctuations on scales $\lesssim 2\,\lambda_{\rm p}$.

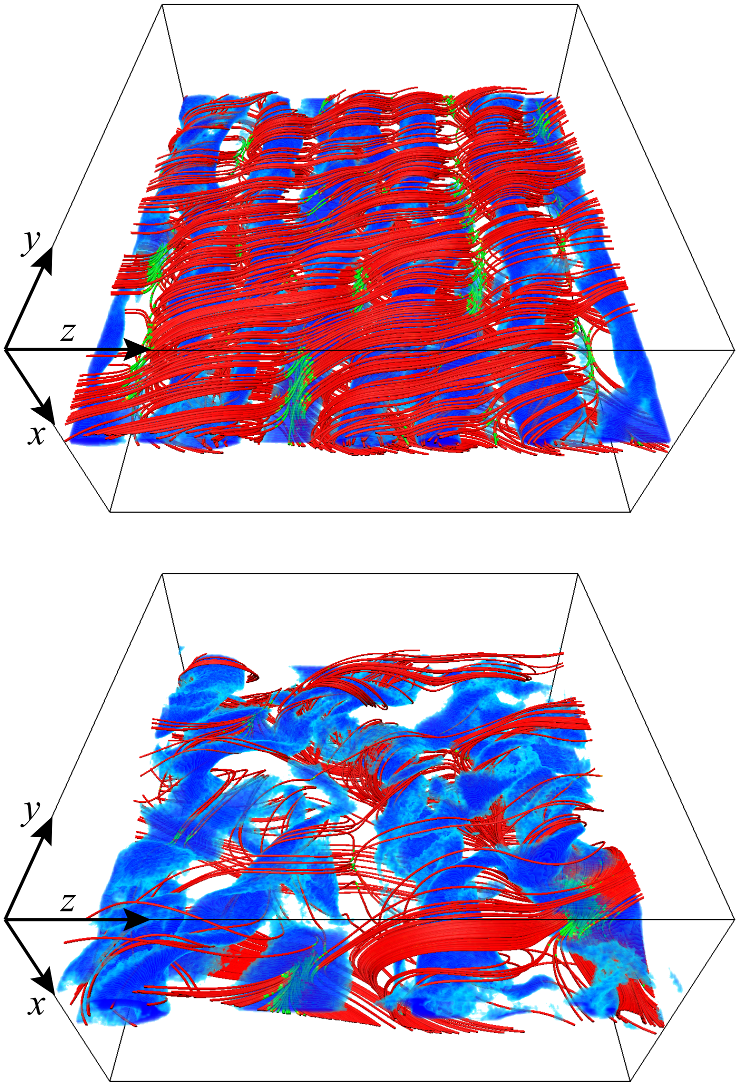

Figure 3.3 A view of the current sheet at $x = L_x/4$ at times $t = 56\,\tau_{c0}$ (upper panel) and $t = 84.5\,\tau_{c0}$ (lower panel). The volume rendering (blue color) shows the region with current density $J > 0.2\,cB_0/(4\pi\lambda_\mathrm{p})$ (upper panel) and $J > 0.12\,cB_0/(4\pi\lambda_\mathrm{p})$ (lower panel). The curves are the magnetic field lines that pass through regions of high parallel electric field $E_{||}$ and thus are actively undergoing reconnection. The fraction of the field line with an enhanced value of $E_{||}$ is rendered in green.

of the current sheets $t = 56\,\tau_{c0}$, during the initial flux rope merging phase, and at $t = 84.5\,\tau_{c0}$, after the flux rope network has become disordered. The field lines shown are those passing through the sites where $E_{||}$ is significant, where the field line color is shown in green and field line geometry is that of an X-line. The reconnected magnetic field snaps back to become incorporated into that of the flux ropes. It is also noticeable that the flux ropes are not cylindrically symmetric but highly flattened (ribbon-like) and exhibit a clear longitudinal twist.

A common measure of magnetic connectivity is the concept of magnetic helicity, but its definition on the entire periodic domain presents unique challenges (e.g., Berger, 1997, and references therein). In spatially and temporally localized, simply-connected reconnection regions, a generalized helicity can be meaningfully defined using the Finn-Antonsen approach (e.g., Berger, 1999). In guide-field reconnection, a change of the generalized helicity implies a change of global field line connectivity (Schindler et al., 1988). Since the generalized helicity is locally generated by the source term $-2\,\mathbf{E}\cdot\mathbf{B} = -2\,\mathbf{R}\cdot\mathbf{B}$ that is closely related to the parallel electric field, the twisting and braiding of the field lines interior to a flux rope is a consequence of (possibly y-coordinate dependent) field connectivity change in the primary X-line region.[4]

3.3.3 Instability Mode Analysis

To identify modes present in the simulation, we carried out Fourier decomposition of the electric and magnetic field for wave vectors $\mathbf{k}$ parallel to the plane of the

---

[4] The sign of $\mathbf{E}\cdot\mathbf{B}$ should be uniform in each individual X-line region.

current sheet, $k_x = 0$, in simulation `S2K025L`. We search for signatures of the tearing and kink instability, as well as for oblique modes, see, e.g., Daughton et al. (2011) and Baalrud et al. (2012). The tearing instability has $\mathbf{k} \parallel \hat{\mathbf{z}}$ and creates a perturbation in B_x and E_y. The kink instability has $\mathbf{k} \parallel \hat{\mathbf{y}}$ and in the linear order creates a perturbation in E_z if a guide field is present. The oblique mode is similar to the tearing mode, except that the wave vector is tilted in the yz plane, that of the initial current sheet. In Figure 3.4, we show Fourier power spectra of the magnetic field component B_x in the linear phase at $t = 19\,\tau_{c0}$ and at the end of the fully nonlinear flux rope merging phase at $t = 37.5\,\tau_{c0}$. We also show the corresponding growth rates evaluated over an interval of $\Delta t = 9\,\tau_{c0}$ preceding each of the two times,

$$\omega_{B_x}(\mathbf{k},t) \equiv \frac{\ln|B_x(\mathbf{k},t)/B_x(\mathbf{k},t-\Delta t)|}{\Delta t}. \tag{3.18}$$

In the figure, the growth rates are multiplied by the light crossing time of the compressed current sheet $\tau_c \approx 0.43\,\tau_{c0}$. The modes with fastest growth rates are clustered around the $k_z = \pm 2\pi(L_z/9)^{-1}$ tearing mode at $t = 19\,\tau_{c0}$, but the fastest growth then shifts to longer wavelengths and oblique directions by $t = 37.5\,\tau_{c0}$.

At the end of the linear phase, the power spectrum for B_x and E_y peaks at $(2\pi)^{-1}(k_y,k_z) = (0, \pm(L_z/9)^{-1})$, which is the tearing mode, consistent with the presence of ~ 9 flux ropes in each current sheet. The peak, however, is broadened, likely by an initial lack of phase coherence on scales given by Equation (3.16). For the tearing mode, $\omega_{\mathrm{KTI}} \sim 0.083\,\tau_c^{-1}$. From this we expect lack of phase coherence on scales $\Delta y, \Delta z \gg \lambda_{\mathrm{coh,KTI}} \sim 5.2\,\lambda_0 \sim 16\lambda_{\mathrm{p}}$, which can explain the observed broadening. However the broadening also allows the possibility that an authentic oblique component is present, too. At the end of the flux rope merging phase, which we define as the time $t \sim (50-60)\,\tau_{c0}$ when the orderly flux rope merging gives way to a more disordered evolution, the peak has shifted to longer wavelengths $(L_z/2)^{-1} \leq (2\pi)^{-1}|k_z| \leq (L_z/6)^{-1}$ with significant additional power occurring in the oblique direction, for $(2\pi)^{-1}k_y \leq (L_y/3)^{-1}$. We emphasize, again, that the oblique power may be an outcome of phase decoherence, as well as of secondary instabilities developing in the nonlinear regime, rather than of a genuine primary oblique instability mode.

Turning to our search for the kink mode, we find that E_z develops nonvanishing albeit weak $\mathbf{k} \parallel \hat{\mathbf{y}}$ power at $t \gtrsim 40\,\tau_{c0}$ with the peak wavelength corresponding to $(2\pi)^{-1}k_y = (L_y/3)^{-1}$. However, most of the power in E_z is still along $\mathbf{k} \parallel \hat{\mathbf{z}}$. Other components of $\mathbf{E}$ and $\mathbf{B}$ are devoid of power along $\mathbf{k} \parallel \hat{\mathbf{y}}$. This suggests that the simulation exhibits no evidence for a linear kink mode and that any variation with y emerges in the nonlinear development of the tearing mode. We compare these results to previous analyses of tearing, kink, and oblique modes in Section 3.4.1 below.

3.3.4 Flux Rope Merger Timescales

Once the tearing instability has produced a collection of nonlinear flux ropes at a time which we will denote with $\tau_{\rm NL}$, the ropes begin to merge. The merging time should scale with the island separation Δz as

$$\tau_{\rm merge} \sim \frac{\Delta z}{\chi v_{\rm A}}, \tag{3.19}$$

where χ is a dimensionless coefficient encapsulating the dynamics of interaction between flux ropes. Then, the number of flux ropes per unit length along the current

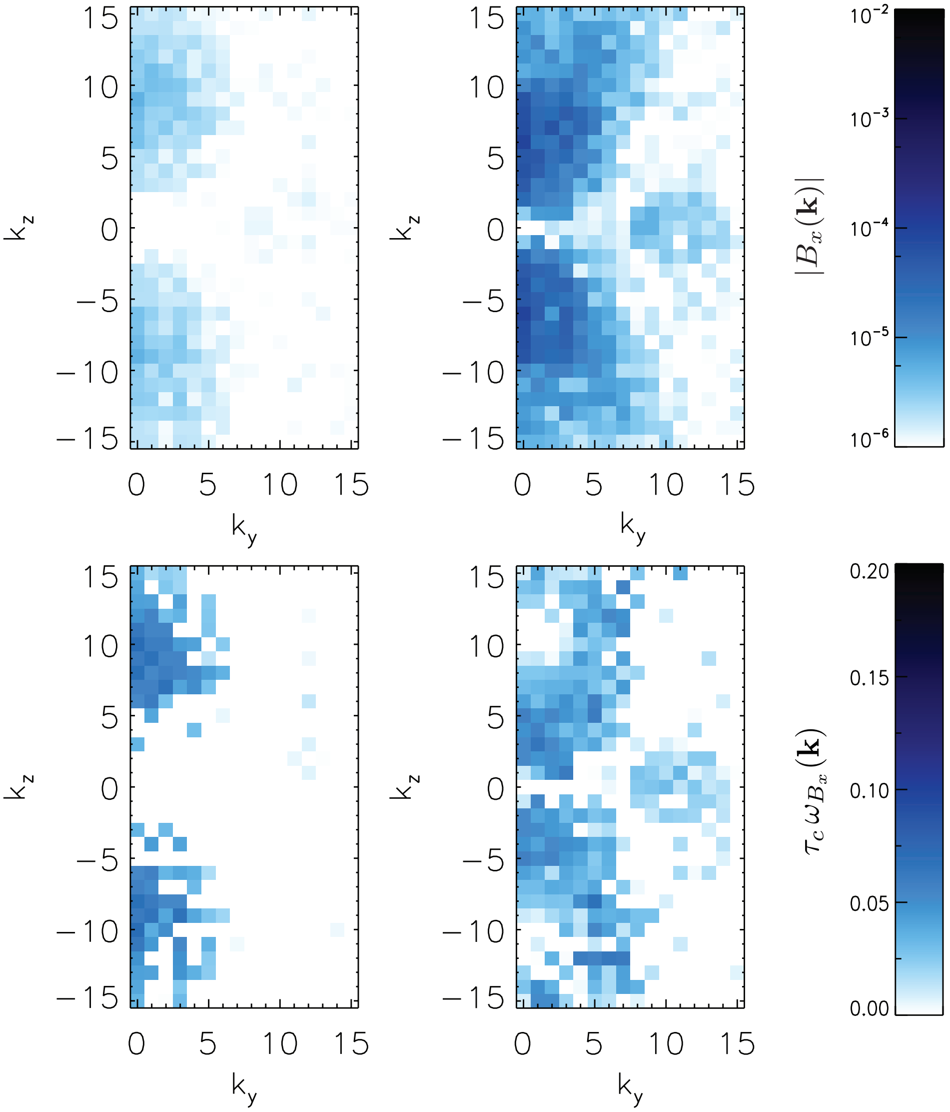

Figure 3.4 Fourier mode amplitudes (top panels) and growth rates (bottom panels) of B_x in one of the current sheets at times $t = 19\,\tau_{c0}$ (left panels) and $t = 37.5\,\tau_{c0}$ (right panels) in simulation `S2K025L`. The amplitude was computed with the Fast Fourier Transform (FFT) via $\mathbf{B}(\mathbf{k}) = (N_x N_y N_z)^{-1} \sum_{i_x=1}^{N_x} \sum_{i_y=1}^{N_y} \sum_{i_z=1}^{N_z} \exp[-2\pi i(i_y k_y/N_y + i_z k_z/N_z)]\,\mathbf{B}(i_x, i_y, i_z)$, where N_x, N_y, and N_z are dimensions of the computational grid, and the initial reversing field $B_0 = 0.1$ in these units. The components of the wave vector are expressed in units of 2π divided by the computational box size along the relevant direction. We only show growth rates for modes with amplitudes $|B_x(\mathbf{k})| > 10^{-6}$.

sheet, which we denote with μ, can be estimated as

$$\mu \sim \frac{1}{\chi v_{\rm A}\,(t-\tau_{\rm NL})}. \tag{3.20}$$

By visually counting the number of flux ropes in each current sheet in simulation `S2K025L` in Figure 3.2, we find that

$$\mu \approx \frac{1.2}{\lambda_{\rm p}}\left(\frac{t}{\tau_{c0}}-16\right)^{-1}, \tag{3.21}$$

suggesting that $\tau_{\rm NL} \approx 16\,\tau_{c0}$ in that simulation. Comparing equations (3.20) and (3.21) we estimate that $\chi \approx 0.4$. Given that the reconnection outflow velocities reach only about a half of the Alfvén velocity in the simulation, it is not surprising that the characteristic velocity associated with flux rope merging $\chi v_{\rm A}$ is only a fraction of the Alfvén velocity.

In reconnection configurations starting from a one-dimensional current sheet lacking any y and z dependence, the three-dimensionality manifested in an emerging y dependence can arise from several effects which include an intrinsic obliqueness of the tearing modes, the development of kink modes, and a lack of coherence on scales on which the phase of the tearing mode is uncorrelated (see Section 3.2.5). In reconnection regimes in which oblique and kink modes are suppressed, flux rope dynamics acquires three dimensional character when the flux rope separation $\sim \mu^{-1}$ starts substantially exceeding the coherence length $\lambda_{\rm coh,KTI}$. From equations (3.14), (3.16), and (3.20), this happens at times

$$t \gg \tau_{\rm 3D,coh} \sim \tau_{\rm NL} + \frac{\lambda}{0.35\,\chi v_{\rm A}\,\beta^{3/2}}, \tag{3.22}$$

yielding an estimate $\tau_{\rm 3D,coh} \sim 36\,\tau_{c0}$ for simulations `S2K025` and `S2K025L`. Indeed, the flux rope geometry in Figure 3.2 seems to acquire three-dimensional character at $t \gtrsim (40-50)\,\tau_{c0}$. This is consistent with the rise of power at oblique wave vectors $k_y > 0$ in Figure 3.4 at similar times. This analysis suggests that the reconnection layers acquire three dimensional structure on relatively short time scales regardless of the growth rates of the oblique and kink modes.

3.3.5 The Reconnection Rate

We now discuss the reconnection efficiency at primary X-lines in the early stages of the evolution of the current sheet. In this regime, the reconnection site can be described using a typical two-dimensional model of reconnection in which two oppositely oriented fields are separated by a thin current layer in which the field reverses. In this layer, magnetic field lines diffuse across the plasma to reconnect at one or more X-lines. Magnetized plasma approaches the central plane of the layer, toward an X-line, with an asymptotic inflow velocity $v_{\rm in}$, which is also known as the reconnection velocity. After passing the X-line, plasma is expelled from the vicinity of the X-line to either side at the outflow velocity $v_{\rm out}$, which is typically assumed to equal the Alfvén velocity $v_{\rm A}$. The orientation of the reconnected magnetic field is approximately perpendicular to that of the original field.

Defining δ and L to be the thickness and length of the current sheet, the conservation of mass from the inflow to the outflow requires

$$\frac{\delta}{L} \approx \frac{v_{\rm in}}{v_{\rm out}} \approx \frac{v_{\rm in}}{v_{\rm A}}. \tag{3.23}$$

The inflow velocity can be related to the electric field in the reconnection region $\mathbf{E}_{\mathrm{RR}}$ which is approximately uniform and equal to the dynamical electric field in the plasma inflow just outside the current sheet

$$\mathbf{E}_{\mathrm{RR}} \approx \frac{1}{c}(\mathbf{v}_{\mathrm{in}} \times \mathbf{B}_0). \tag{3.24}$$

Because the inflow velocity and the reversing magnetic field are perpendicular, the dimensionless reconnection rate r_{rec}, defined as the ratio of the inflow and the outflow velocity, and the electric field in the reconnection region, can be related via

$$r_{\mathrm{rec}} \equiv \frac{v_{\mathrm{in}}}{v_{\mathrm{out}}} = \frac{E_{\mathrm{RR}}}{(v_{\mathrm{A}}/c)B_0}. \tag{3.25}$$

This calculation of the reconnection rate makes the common assumption that the outflow velocity is equal to the Alfvén velocity in the inflow region. In our simulations, this assumption is not valid as the outflow velocity is smaller than the Alfvén velocity.

In simulation `S2K025`, we find that indeed, the total electric field in the reconnection region is approximately uniform across the current sheet. Given that the Alfvén velocity is $v_{\mathrm{A}} \approx 0.7\,c$ in this simulation, the dimensionless reconnection rate calculated from the electric field following Equation (3.25) is $r_{\mathrm{rec}} \approx 0.05$ at $t = 37.5\,\tau_{c0}$, in the middle of the initial flux rope merging phase. This is similar to the peak reconnection rates of about $0.07-0.1$ found in other three dimensional PIC simulations with higher magnetizations (e.g., Zenitani & Hoshino, 2008; Liu et al., 2011). This comparison assumes that reconnection at the largest X-line can be described accurately by a two dimensional reconnection model. The inflow and outflow velocities at $t = 37.5\,\tau_{c0}$ are $v_{\rm in} \approx 0.02\,c$ and $v_{\rm out} \approx 0.27\,c$, respectively. The peak plasma outflow velocity slightly later in the simulation reaches $v_{\rm max} \approx 0.39\,c$, just above half of the Alfvén velocity. Since the Alfvén velocity is not ultrarelativistic, it is not surprising that we are not finding ultrarelativistic flows in any of our simulations.[5] If the dimensionless reconnection rate is calculated directly from the inflow and outflow velocities, it is $r_{\rm rec} \approx 0.08$. The difference between the outflow velocity and the Alfvén velocity, which are equal in idealized two dimensional reconnection models, is likely due to the dynamical, non-steady state nature of the plasma flow in our simulations; similar discrepancies were found by Bessho & Bhattacharjee (2010) in their nonrelativistic PIC simulations, and Takahashi et al. (2011) in their relativistic MHD simulations.

3.3.6 Generalized Ohm's Law Analysis

A central question in magnetic reconnection is the nature of the process that facilitates the violation of flux freezing in the diffusion region. The total electric field **E** in a collisionless plasma is related to the magnetic field **B** and to moments of the particle distribution function of particle species *s* by the generalized Ohm's law (e.g., Krall & Trivelpiece, 1973)

$$\begin{aligned}\mathbf{E} + \frac{1}{c}\langle \mathbf{v}_s \rangle \times \mathbf{B} \quad = \quad & \frac{1}{q_s n_s} \nabla \cdot \mathbf{P}_s \\ & + \frac{1}{q_s} \left(\frac{\partial \langle \mathbf{p}_s \rangle}{\partial t} + \langle \mathbf{v}_s \rangle \cdot \nabla \langle \mathbf{p}_s \rangle \right), \end{aligned} \tag{3.26}$$

---

[5] To obtain an ultrarelativistic flow, one must have $\gamma_{\rm A} \equiv [1-(v_{\rm A}/c)^2]^{-1/2} \gg 1$, but this may not be sufficient, e.g., Zenitani & Hoshino (2008) found maximum flow speeds of $v_{\rm out} \approx (0.7-0.8)\,c$ even in their two-dimensional simulations with high magnetization and no guide field.

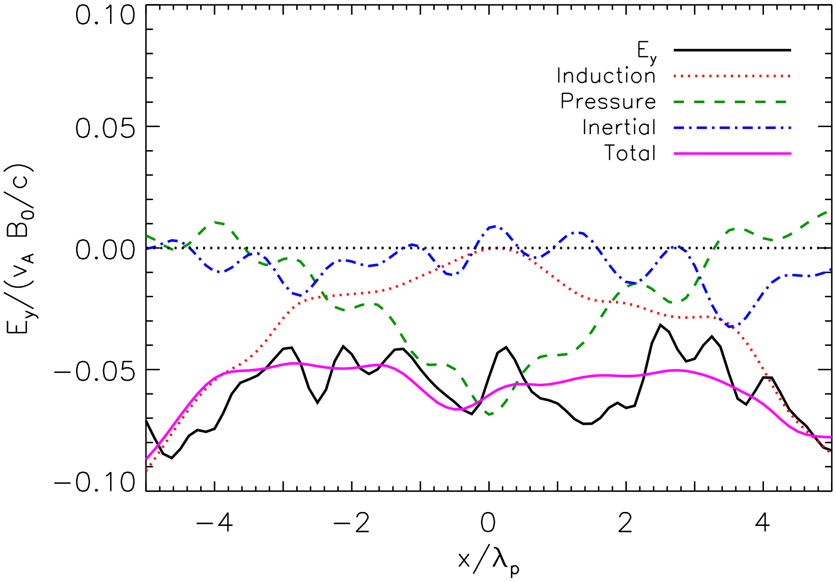

Figure 3.5 The variation of the terms in the generalized Ohm's law of the electric field across the current sheet near the largest X-line at $t = 47\,\tau_{c0}$. The center of the current sheet is at $x = 0$. The lines represent the measured electric field E_y (solid, black), the induction component (dotted, red), the pressure component (dashed, green), the inertial component (dot-dashed, blue), and the total (pink, solid) of the three calculated electric field terms. The generalized Ohm's law mandates that the latter should equal the actual electric field, which is the case apart from fluctuations arising from numerical discreteness. The Ohm's law components are smoothed with a Gaussian kernel of standard deviation $0.33\lambda_\mathrm{p}$. It is clear that at the X-line, the pressure term dominates over the inertial term.

where $\langle \mathbf{v}_s \rangle$ and $\langle \mathbf{p}_s \rangle$ are the average velocities and momenta of particles, $\mathbf{P}_s$ is the pressure tensor, n_s is the particle number density, and q_s and m_s are the particle charge and mass. Averaging over the two species $s = \{i, e\}$ we recover the non-ideal, reconnection electric field $\mathbf{R}$ on the left hand side of Equation (3.26). In collisionless plasmas with long-lived magnetization (counterexamples being the plasmas excited by kinetic instabilities, such as the filamentation instability), the pressure and the inertial terms on the right hand side of Equation (3.26) are small almost everywhere. Both of these terms, however, can become important in the diffusion region of a magnetic reconnection layer.

To determine which terms give rise to $\mathbf{R}$, we calculate the terms on both sides of Equation (3.26) as they vary across the X-line in simulation `S2K025`. We find no significant reconnection electric field in the x or z directions; there is only a reconnection field in the y direction, consistent with two dimensional models of reconnection. Meanwhile, during the linear growth phase of the tearing instability, spatial gradients are present only in the x and z directions, implying that only the off-diagonal xy and zy components of the pressure tensor contribute to the first term on the right hand side of Equation (3.26). In particular, we find that most of the pressure term is provided by $\partial P_{xy}/\partial x$. Figure 3.5 shows the y component of various terms contributing to the electric field. The pressure term is dominant, and the total of the inductive, pressure, and inertial terms approximates the electric field, with departures arising from shot noise.

These results are consistent with previous investigations (e.g., Bessho & Bhattacharjee, 2005, 2007; Schmitz & Grauer, 2006) that found that the spatial variation of off-diagonal terms of the pressure tensor was the main contributor to the reconnection electric field. The results of our Ohm's law calculation with a moderate guide field, combined with the similar results of other simulations with strong guide field (Che et al., 2011, nonrelativistic simulations) and no guide field (Liu et al., 2011, relativistic simulations), constitute evidence that for moderate magnetizations and all guide field magnitudes, the reconnection electric field is produced by the pressure term in three as well as in two dimensions. In an exception to this agreement, Hesse & Zenitani (2007) found that the inertial term became important in highly magnetized relativistic current sheets with a guide field, which is a regime we do not probe.

3.3.7 Particle Energization

We now turn to the topic of magnetic to kinetic energy conversion and particle energization. Figure 3.6 shows the evolution of the total magnetic field, electric field, and particle kinetic energies in simulations `S2K025` and `S2K025L`. Energy was conserved to within 0.3% in all the simulations. The initial kinetic and electromagnetic energy fraction differs in the two simulations because the initial current sheet occupies a fraction of the volume of the simulation box that is twice as large in simulation `S2K025` as in `S2K025L`. The trends, however, are consistent taking the volume difference into account. An initial reversible fluctuation of the magnetic-to-kinetic ratio during the first $\sim 20\,\tau_{c0}$ is associated with the readjustment to pressure equilibrium. Dissipative conversion of magnetic to kinetic energy seems to first occur in a couple of bursts at $t \sim (30-40)\,\tau_{c0}$ and $t \sim (50-60)\,\tau_{c0}$. Then, the

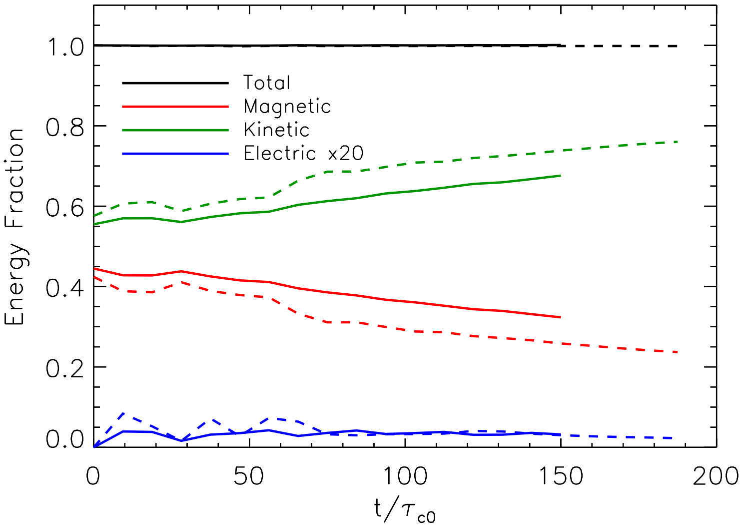

Figure 3.6 Evolution of the total magnetic energy (red lines), particle kinetic energy (green lines), and the energy in the electric field multiplied by 20 for clarity (blue lines), all normalized to the total initial energy, in runs `S2K025L` (solid lines) and `S2K025` (dashed lines). The time is expressed in units of the light crossing time of the current sheet τ_{c0}.

system enters a phase, lasting until the end of the simulation, in which dissipation is relatively steady. The two bursts seem to coincide with initial tearing instability and the formation of ~ 9 flux ropes (magnetic islands) across the simulation box per current sheet, and the subsequent flux rope merging, yielding ~ 3 flux ropes per current sheet by $t \sim 60\,\tau_{c0}$, after which the overall geometry of the reconnection layer becomes fully three dimensional. By $t = 150\,\tau_{c0}$, about 40% and 20% of the initial magnetic energy is converted into particle kinetic energy in the smaller and larger simulation, respectively.

In Table 3.1, we show $|\Delta\mathcal{E}_B|/\mathcal{E}_B$, the fraction of magnetic field energy that is converted to particle kinetic energy in each simulation. Here, we define $\mathcal{E}_B$ to be the initial energy in the total magnetic field, which includes both the reversing field and the guide field, and $\Delta\mathcal{E}_B$ is the change of the total magnetic energy from the beginning to the end of the simulation. The energy in the reversing field only is a factor of $\approx (1+\kappa^2)^{-1}$ smaller than $\mathcal{E}_B$ if we ignore the reduction of the reversing field in the current sheets. Independent of this correction, $|\Delta\mathcal{E}_B|/\mathcal{E}_B$ exhibits a strong decreasing trend with increasing guide field strength κ. It also exhibits a weaker increasing trend with σ; note that the very high $|\Delta\mathcal{E}_B|/\mathcal{E}_B$ in simulation `S1K0` is the result of destructive current sheet interaction. In the simulations with $\kappa \geq 0.5$, and also in simulation `S1K025` with $\kappa = 0.25$ and $\sigma = 1$, the conversion of magnetic to kinetic energy is substantially weaker than in the other simulations. We also note that the rate of development and evolution of the flux rope network exhibits similar behavior in the sense that flux ropes of a given size assemble later ($\sim 80\,\tau_{c0}$ vs. $\sim 40\,\tau_{c0}$) in the $\kappa \geq 0.5$ and `S1K025` simulations.

The variation of $|\Delta\mathcal{E}_B|/\mathcal{E}_B$ and of the overall evolution rate with κ is likely a consequence of the decrease of plasma compressibility with increasing guide field strength (e.g., Zenitani & Hesse, 2008, and references therein). Since most of the magnetic to particle kinetic energy conversion takes place during the relaxation of reconnection field lines to their final equilibrium in flux ropes, the less compressible case of a strong guide field leads to larger flux ropes and smaller energy conversion.

3.3.7.1 Energization Efficiencies

Figure 3.7, top panel, shows the energy spectrum at a reference time $t = 150\,\tau_{c0}$ in six of the smaller size simulations. Each of the spectra exhibits a peak in agreement with the initial thermal distribution as well as a new tail resulting from particle energization extending to maximum Lorentz factors in the range $\gamma_{\max} \sim 30-50$ (for the precise values at the end of the simulations, see Table 3.1). The simulations with the strongest particle energization are those with higher magnetization $\sigma = 2$ and zero or moderate guide field $\kappa \leq 0.25$. Rather weak energization is seen in all simulations with strong guide field $\kappa = 1$ and in the simulation `S1K025` with weak magnetization and moderate guide field. We observe an intermediate level of energization in simulation `S1K0` with weak magnetization and no guide field; this simulation is unique in that at the reference time, the two current sheets have already interacted with each other. It is important to note that all of these spectra are calculated at the same reference time; therefore, these differences in energization efficiency might result from the differing energy conversion rates discussed in the previous section. However, we will find below that the spectrum

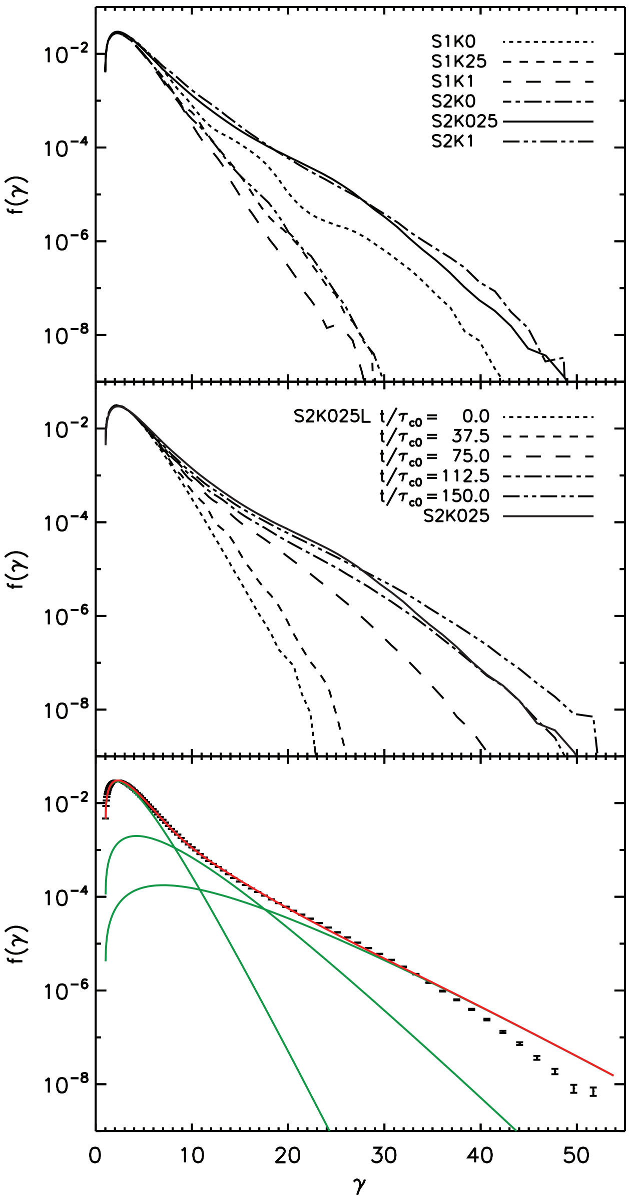

Figure 3.7 Particle energy spectra in the six smaller size simulations at time $t = 150\,\tau_{c0}$ (top panel; see legend) and in the large size simulation `S2K025L` at five different times (middle plan; see legend) where for reference we include the $t = 150\,\tau_{c0}$ spectrum from the corresponding smaller size simulation `S2K025` (solid line). The bottom panel compares the final spectrum to a a model containing three thermal populations at different temperatures (red line; see text), including the spectra of the three individual thermal components (green lines). The spectra and uncertainties are calculated from a random sample containing 5% of the particles in the simulation.

exhibits the strongest evolution during the flux rope merging phase; because the merging phase is complete by the reference time in all the simulations, it seems that the differences in the level of energization between the runs cannot be attributed solely to the differences in the rate of reconnection.

In order to quantify the degree of particle energization in the simulations, following Zenitani & Hoshino (2007), we compute the particle kinetic energy K_{ener} contained in the difference between the measured spectrum and a thermal spectrum containing the same number of particles and the same energy at particle energies at which the measured spectrum is in excess of the thermal spectrum. We then calculate the ratio of K_{ener} to the total particle kinetic energy K, both computed at the end of the simulation. It is worth remarking that we do *not* go as far as Zenitani & Hoshino (2007, 2008) to identify the particles contributing to K_{ener} as a nonthermal population; the following section we show this tail may be better described with a combination of several thermal populations. Consistent with the trend observed in the shape of the spectrum, we find that K_{ener}/K, which is shown in Table 3.1, is the largest in the simulations with the higher magnetization $\sigma = 2$ and at most moderate guide field $\kappa \leq 0.25$ and is the smallest in simulation `S1K1` with lower magnetization $\sigma = 1$ and strong guide field $\kappa = 1$. Thus, in contrast with Zenitani & Hoshino (2008), we find that K_{ener}/K decreases with increasing guide field. The reason for this dissimilarity seems to be that the kink instability does not compromise the development of reconnection in any of our simulations other than `S1K0`; even in that run, a minor flux rope merger occurs before the current sheet is disrupted. Without the kink instability, more particle energization can take place.

Another possibly more interesting measure of particle energization is the ratio of the kinetic energy in energized particles to the magnetic energy converted to kinetic energy. Combining equations (3.1) and (3.6) with the definition of Γ, and taking account of the conservation of energy $K(t) = K(0) + |\Delta\mathcal{E}_B|$, this ratio can be written as

$$\frac{K_{\text{ener}}}{|\Delta\mathcal{E}_B|} = \left(\xi \left| \frac{\Delta\mathcal{E}_B}{\mathcal{E}_B} \right|^{-1} + 1 \right) \frac{K_{\text{ener}}}{K}. \tag{3.27}$$

Here, the dimensionless coefficient ξ is the ratio of the initial particle kinetic energy density to the magnetic energy density including the guide field, given by

$$\xi \approx \frac{1}{\sigma(1+\kappa^2)} \left(\frac{1}{\Gamma - 1} - \frac{1}{\theta_0} \right), \tag{3.28}$$

where $\theta_0 \equiv T_0/(m_e c^2)$ and all the quantities entering the definition of ξ are evaluated at the beginning of the simulation. The ratio, which ranges between 11% and 38%, is shown in Table 3.1. The trend seen at low magnetization $\sigma = 1$ is an increase of $K_{\text{ener}}/|\Delta\mathcal{E}_B|$ with guide field strength, approximately the opposite of that seen in K_{ener}/K. The trend can be understood by noting that, as we will find in Section 3.3.7.3, most of the particle energization contributing to K_{ener} occurs in X-line regions. Meanwhile, $\Delta\mathcal{E}_B$ measures the change in magnetic energy both during reconnection in X-line regions and, more importantly, during the subsequent contraction of reconnected magnetic field lines into flux ropes. Therefore the ratio $K_{\text{ener}}/|\Delta\mathcal{E}_B|$ should indeed increase with stronger guide field because the guide field reduces plasma compressibility and moderates field line contraction.

3.3.7.2 Energy Spectrum Evolution and Structure

Figure 3.7, middle panel, shows the evolution of the particle energy spectrum in the large simulation `S2K025L`. The spectrum evolves little during the linear tearing phase and the first round of flux rope merging at $t < 37.5\,\tau_{c0}$, but then it quickly hardens by $t = 75\,\tau_{c0}$ as the flux rope merging concludes and the reconnection layer transitions to a disordered state. Further hardening takes place until the end of the simulation at $t = 150\,\tau_{c0}$. For comparison, we also show the particle energy spectrum at this time from the smaller, equivalent simulation `S2K025`. The two spectra are consistent at Lorentz factors $\gamma \lesssim 28$ indicating convergence with increasing box size at relatively low energies. However, the larger simulation has progressively more particles at still higher energies, with Lorentz factors reaching $\gamma \approx 50$.

We proceed to model the entire spectrum in simulation `S2K025L` as is, while keeping in mind that in a still larger simulation, further evolution of the spectrum at the highest energies is likely to be expected. We experimented with composite populations containing thermal as well as power law components, and found that at Lorentz factors $\gamma \lesssim 35$, a model containing three thermal populations, each described by a relativistic Maxwellian, seems to work best. The model and the three thermal components are shown the bottom panel of Figure 3.7. We fix the temperature of the first population to be equal to the temperature of the initial plasma, $T_1 = m_e c^2$. A least-squares fit yielded temperatures $T_2 = 2.1\,m_e c^2$ and $T_3 = 3.5\,m_e c^2$ for the second and third population, respectively. The first component contains 85% of the particles and 70% of particle kinetic energy, the second component contains 13% of particles and 24% of particle kinetic energy, and the third component contains only 2% of the particles and 6% of particle kinetic energy. We show in Section 3.3.7.3 that the energized populations and especially the particles with $\gamma > 30$ are located close to the primary X-lines (within about one Larmor radius) and in flux ropes.

3.3.7.3 Energization Sites and Mechanism

To gain insight in the nature of the particle energization process, in simulation `S2K025` we select a number, $N_{\text{trace}} = 1841$, of particles reaching the highest energies, corresponding to Lorentz factors $\gamma \geq 32$, at the end of the flux rope merging phase at $t = 75\,\tau_c$. We trace the orbits of these particles throughout the merging phase over the time interval $47\,\tau_c \leq t \leq 75\,\tau_c$. This is the period during which the particles experience coherent energization. The particles generally begin in the current sheet and have momenta that, on average, are aligned with the direction of the electric force $q_j\mathbf{E}$ in the current sheet, with a median inclination of $\approx 30^\circ$ from the direction of the force.

During flux rope merging, the traced particle energies increase approximately linearly. However at the instance of merging for the two largest flux ropes, which occurs at $t \approx 61\,\tau_c$, most of the particles incur an energy increment accounting for $\approx 15\%-20\%$ of the total energy gain. Following the flux rope merging phase, some traced particles gain additional energy, but others lose energy, and both the gain and the loss could be considered manifestations of a thermalization process. Indeed, the energization K_{ener}/K does not further increase after the flux rope merging is complete.

To pin down the geometric location of particle energization, for the traced

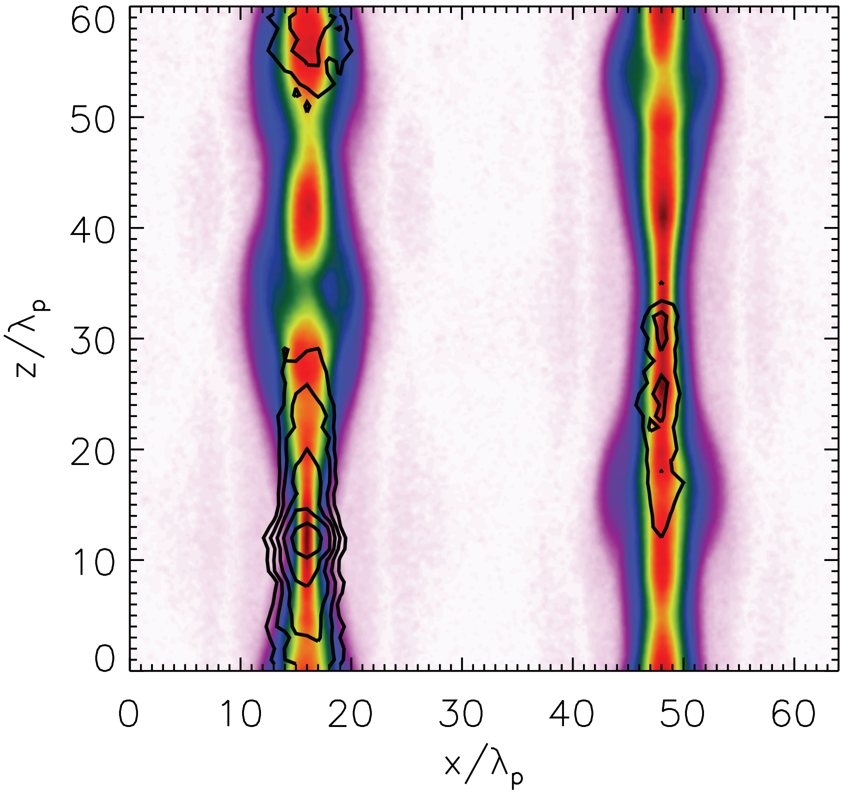

Figure 3.8 The locus of particle acceleration in simulation `S2K025` in the interval $47\,\tau_c \leq t \leq 70\,\tau_c$ corresponding to the primary particle energization phase. The contours show the logarithm of the y-averaged energy gain per particle per unit volume $\Upsilon(x,z)$ for the ≈ 2000 highest energy particles in the simulation in this interval (see text); the lowest contour corresponds to 0.04 times the peak value of Υ. The underlying color image shows the projection in the xz plane of the current density J_y averaged over the interval; the current density is scaled linearly from light purple to dark red.

particles, we compute the total energy gain per particle per unit volume averaged over y via

$$\begin{aligned} \Upsilon(x,z) &= \frac{1}{N_{\rm trace}} \sum_{j=1}^{N_{\rm trace}} \frac{q_j}{m_e c^2} \int \mathbf{v}_j(t) \cdot \mathbf{E}[\mathbf{x}_j(t)] \\ &\quad \times \delta[\mathbf{x} - \mathbf{x}_j(t)]\, dt\, \frac{dy}{L_y}, \end{aligned} \tag{3.29}$$

where $\mathbf{x}_i(t)$ and $\mathbf{v}_i(t)$ are the position and velocity of traced particle j at time t, δ is the Dirac delta, and the time integral covers the flux rope merging phase. In Figure 3.8, we compare the energy gain Υ to the corresponding absolute value of the y- and t-averaged current density J_y. The highest energy gain is clearly associated with the largest, best defined X-line on the lower left. Substantial energization is also detected in the outflow regions flanking this X-line. The total energy gain in the flanking regions, seen vertically above and below the X-line in the figure, is larger than that in the narrow vicinity of the X-line.

It is puzzling that the overall energization efficiency is weaker in the second current sheet shown on the right in the figure. In both sheets, the energization seems to be associated with some current density maxima but not with others. Energization seems to prefer long, continuous X-line regions with thin current sheets, perhaps because these are the sites of plasma inflow.

In Figure 3.9 we show the orbits of six representative traced particles belong to the current sheet showing on the left of Figure 3.8. They all start in the current sheet within $\sim 8\,\lambda_{\rm p}$ from the X-line and have initial Lorentz factors in the range $\gamma \sim 16-23$. While in the X-line region, the particles oscillate across the current sheet on Speiser-like orbits. The particles reach Lorentz factors of $\gamma \sim 32$ (this is how they were selected; see Figure 3.9, right panel) after drifting from the X-line region into the neighboring islands. At the final Lorentz factor, the particles have Larmor radii evaluated using the reversing magnetic field $r_{\rm L} = \gamma m_e c^2/(eB_0) \sim 17\,\lambda_{\rm p}$, larger than the width of the flux ropes in the simulation. Because the magnetic field in flux ropes is only somewhat enhanced compared to the reversing field, the accelerated particles are not easily trapped within the flux ropes.

The electric field is nearly uniform across the reconnection region, hence the electric force acting on a particle trapped in a Speiser orbit is approximately constant in time provided that the the inclination of the particle's velocity relative to the direction of the electric force is relatively small. Most of the traced particles happen to fulfill this condition. Then, the work done by the electric force on the particle is independent of the initial particle energy. We have also checked that, we expect for the Speiser orbit acceleration, it is the y-component of the electric field that contributes the most to the energy gain. The average total energy gain per particle obtained by integrating Υ over the volume of the simulation is $\Delta\gamma \approx 12$.

Because the rate of particle acceleration in the y-direction is constant in time, the total energy gain is limited by the time the particle spends in the X-line region (see, e.g., Cerutti et al., 2012a). Particles displaced from the very center of the X-line experience a Lorentz force due to the reconnected magnetic field B_x that deflects them away from the X-line and toward the flux rope. Once a particle has left the X-line region and entered a flux rope, it no longer experiences coherent acceleration.

While Drake et al. (2006) found that acceleration internal to flux ropes

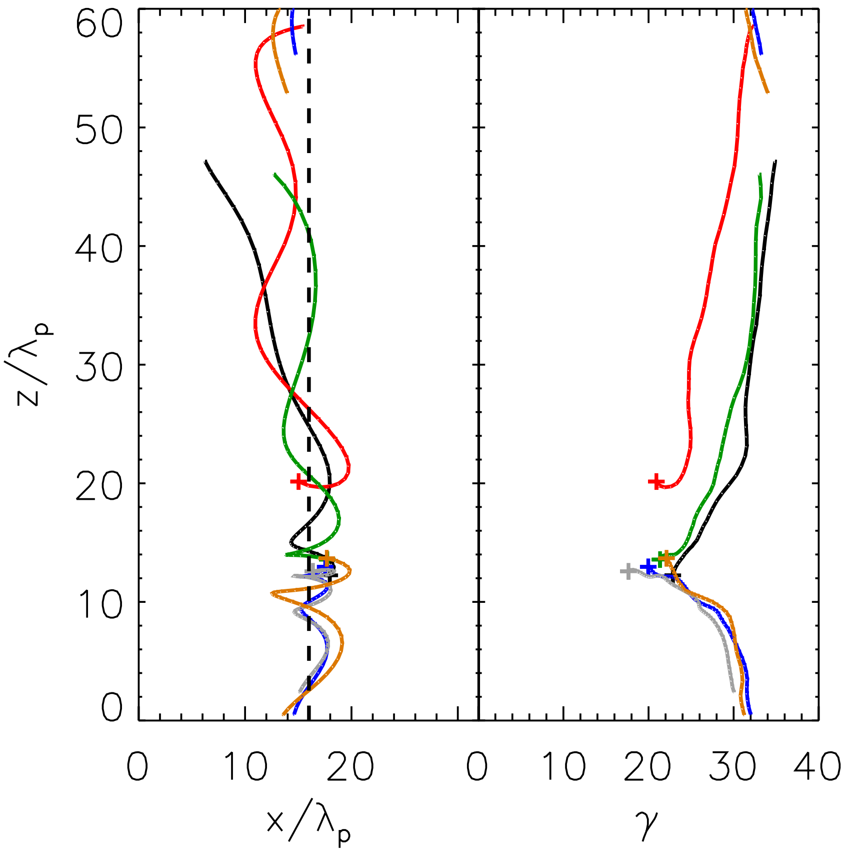

Figure 3.9 The trajectories of six representative particles that attain Lorentz factors $\gamma > 30$ in the simulation `S2K025` (left panel), and the evolution of the Lorentz factor as a function of coordinate z (right panel). The dashed line indicates the center of the current sheet. Note that many of the particles start at $z \approx 12.5\,\lambda_\mathrm{p}$ and close to the center of the current sheet; this is the location of the large X-line on the lower left in Figure 3.8. The crosses indicate the particle position at the beginning of the acceleration phase.

occurred when guide field was present, this mechanism has not typically been important in other works for the guide field case (Fu et al., 2006; Huang et al., 2010). We note that unlike the particles traced in Drake et al. (2006), our energized particles already have very large Larmor radii by the time they transition from the X-line region into a flux rope and the magnetic geometry and dynamics also seem markedly different.

3.3.7.4 Angular Distribution of The Highest Energy Particles

If the motion of high energy particles accelerated in a reconnection region is anisotropic, then the synchrotron radiation emitted by these particles carries angular dependence. Radiation emitted by beams of accelerated particles can come in and out of view of an observer; at any given time, the observer detects only the radiation emitted by particles with momenta making angles $\lesssim\gamma^{-1}$ from the line of sight. Such beaming could explain the temporal variability often seen in astrophysical sources, such as GRBs (e.g., Zhang & Yan, 2011; McKinney & Uzdensky, 2012) and blazars (e.g., Nalewajko et al., 2011), that may be powered by reconnection. It could also influence the characteristics of the gamma-ray emission from other candidate reconnection-powered astrophysical sources, such as magnetars (Thompson & Duncan, 1995; Lyutikov, 2003; Parfrey et al., 2012) and pulsar wind nebulae (e.g., Lyubarsky & Kirk, 2001; Sironi & Spitkovsky, 2011; Cerutti et al., 2012a). We investigate the angular distribution of the momenta of the highest energy particles in our simulation, and compare with the results of Cerutti et al. (2012b), who recently reported a high degree of beaming in a two-dimensional PIC simulation of pair plasma reconnection initialized with $\sigma \sim 40$.

Figure 3.10 shows the orientations of the momenta of all of the electrons and ions with $\gamma \geq 30$ in simulation `S2K025L`. We select the larger simulation for this analysis to include a larger range of three-dimensional effects to which the particle anisotropy may be sensitive. The median inclination of electrons (ions) to the y-axis (negative y-axis) is $\sim 30^\circ - 40^\circ$, hence a half of particles in each charge species occupies a fraction $\Delta\Omega/(4\pi) \sim 0.06 - 0.12$ of the full solid angle. This moderate degree of beaming is a natural consequence of X-line acceleration by the electric field in the reconnection region, since the electric field is uniform and accelerates the particles preferentially along the y-axis.

The degree of beaming, however, is much smaller than in Cerutti et al. (2012b), where the highest energy particles, with Lorentz factors $\gamma > 40$, were strongly beamed, occupying a solid angle fraction as small as $\Delta\Omega/(4\pi) \sim 0.01$. The dependence of the beaming on the parameters of the reconnection layer can crudely be understood as follows. Assuming for simplicity that the plasma is ultrarelativistic and highly magnetized, and that the energy of the accelerated particle increases by a large factor, the components of its momentum following acceleration are approximately $p_y \sim \Delta\gamma m_e c$ and $p_x \sim p_z \sim \gamma_{\text{init}} m_e c$, where γ_{init} is the initial Lorentz factor of an accelerated particle, and $\Delta\gamma$ is the Lorentz factor gain during acceleration, which increases the particle momentum in the y direction. Therefore, the particle's degree of beaming is proportional to $\Omega/(4\pi) \propto (p_x^2 + p_z^2)/p_y^2 \propto (\Delta\gamma/\gamma_{\text{init}})^{-2}$. Using $\Delta\gamma \propto eE\Delta t/(m_e c)$, where $E \sim r_{\text{rec}} v_{\text{A}} B_0/c$ is the accelerating electric field (Equation 3.25) and $\Delta t \sim \tau_{\text{merge}} \sim \Delta z/v_{\text{A}}$ is the duration of acceleration (Equation

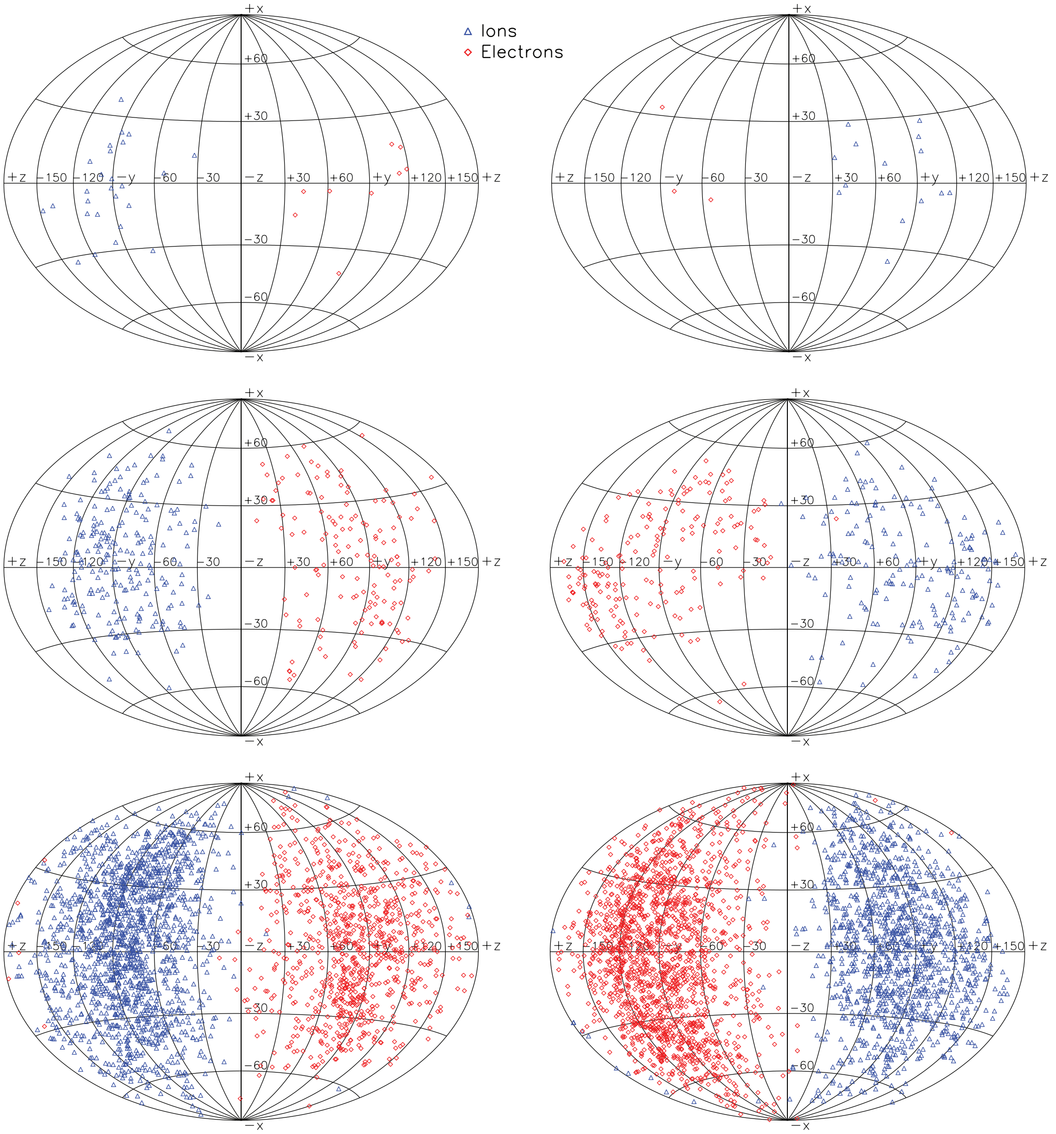

Figure 3.10 Angular distribution of the momenta of particles with Lorentz factors $\gamma \geq 30$ at times $t = 56\,\tau_{c0}$ (top panels), $t = 66\,\tau_{c0}$ (middle panels), and $t = 84\,\tau_{c0}$ (bottom panels) in simulation `S2K025L` shown in the Aitoff projection. The panels of the left and right show the particles near the current sheets at $x = L_x/4$ and $x = 3L_x/4$, respectively. The plotted particles were selected from a random sample containing 5% of the particles in the simulation.

3.19), and expressing the magnetic field in terms of the magnetization parameter $B_0^2 \propto \sigma n_0 T_0 \propto \sigma \langle\gamma\rangle \theta_0 (m_e c^2)^2/(e\lambda_{\rm p})^2$, where $\theta_0 \equiv T_0/(m_e c^2)$ (Equations 3.6 and 3.7), we obtain

$$\frac{\Omega}{4\pi} \propto \frac{1}{\sigma}\left(\frac{\gamma_{\rm init}^2}{\theta_0 \langle\gamma\rangle}\right)\left(r_{\rm rec}\frac{\Delta z}{\lambda_{\rm p}}\right)^{-2}. \tag{3.30}$$

This result is accurate in the regime $\theta_0 \gtrsim 1$ whereas Cerutti et al. set up their simulation with $\theta_0 = 0.15$. Ignoring this concern, we estimate that the factors in parentheses on the right hand side of Equation (3.30) are similar in our simulation and in Cerutti et al. We therefore expect that the beaming solid angle $\Omega/(4\pi)$ in their simulation should be smaller by the inverse ratio of the magnetizations in the two simulations, which is $(\sigma_{\rm Cerutti}/\sigma_{\tt S2K025L})^{-1} = 0.05$. Thus, the crude expectation is that the particles accelerated in Cerutti et al. should more beamed within a solid angle over an order of magnitude smaller than in our simulation, consistent with the observed beaming ratio between the two simulations. In conclusion, this analysis suggests that high magnetization and large physical size of the X-line region can both give rise to a high degree of beaming.

3.4 Discussion

3.4.1 Tearing, Kink, and Oblique Modes

A number of studies have examined the three-dimensional evolution of a current sheet without a guide field; here, we focus on results relevant to pair plasmas. Zenitani & Hoshino (2008) found that the kink mode dominated the tearing mode in antiparallel reconnection. In contrast, Liu et al. (2011) found that the initial evolution was dominated by the tearing mode, leading to an initial merging phase similar to that in two dimensions. After the completion of the merging phase, however, they found that a secondary kink instability set in and rendered the reconnection region turbulent and filled with plasmoids covering a range of dimensions. The reduced influence of the kink mode in the simulations of Liu et al. (2011) may be explained by initially relativistic drift velocities of $0.82\,c$. Analytical calculations suggest that while the kink mode dominates at low drift velocities, the tearing mode becomes dominant for drift velocities exceeding $0.6\,c$ (Zenitani & Hoshino, 2007). However, the growth rate of the kink mode, unlike the tearing mode, depends strongly on the initial structure of the current sheet (Daughton, 1999).

In guide field reconnection, previous three dimensional studies have found that the kink mode is stabilized, but, under certain conditions, oblique modes dominate the evolution. Zenitani & Hoshino (2008) found that with a guide field present, the tearing mode and a sausage-like mode are combined into an oblique "relativistic-drift-sausage-tearing" mode. Daughton et al. (2011) found that in electron-ion plasmas, oblique modes give rise to a network of interconnected flux ropes. In both cases, the nonlinear development of the oblique modes led to turbulence in the current sheet. In nonrelativistic electron-ion simulations with a strong guide field, Che et al. (2011) found that at low temperatures, the current sheet filamented and that the turbulent transport played a key role in the diffusion of the field. At higher, albeit still nonrelativistic temperatures, their filamentation instability became weaker and the current sheet did not become turbulent.

One must be aware of the possibility that unlike the tearing mode, the other instabilities identified in simulations starting with Harris sheet equilibria are idiosyncrasies of the initial configuration and have little to do with astrophysical magnetic reconnection. Our simulation, starting from a non-Harris like configuration lacking a density enhancement in the initial current sheet, should help discern artifacts of the initial configurations from genuine properties of reconnection dynamics.

We have found that in our simulations, in which we considered guide fields at most equal to the reversing field, the tearing mode dominates the evolution at all guide field strengths except in Run `S1K0`. To understand the difference between our simulations without guide field and the studies that detected a dominant kink mode, we have carried out two dimensional simulations with $\sigma = 2$ and $\kappa = 0$ in both the xy and xz planes. We find that the kink mode grows faster in a Harris current sheet, while the tearing mode grows faster starting with our initial conditions. Our initial configuration seems to enhance the growth rate of the tearing mode and inhibit the growth of the kink mode. The plasma inflow occurring during readjustment to equilibrium may enhance growth of the tearing mode above the value expected in pressure equilibrium. Furthermore, our initial configuration may be endowed with a somewhat stronger drift velocity shear than the Harris sheet and this could stabilize the kink mode (Volponi et al., 2000) while having little effect on the tearing mode (Roytershteyn & Daughton, 2008).

In our simulations with guide field, we do not find any oblique modes similar to those observed by Zenitani & Hoshino (2008) or Daughton et al. (2011). Therefore, the only linear mode present is the tearing mode; the lack of phase coherence is the mechanism responsible for breaking translational symmetry. Because this mechanism is a general result of causal constraints, it is likely that this mechanism

Table 3.2. Tearing Instability Parameters

Run	β_0	f_c [a]	$N_{\rm FR}$ [b]	$\tau_{c0}\omega_{\rm KTI,t}$ [c]	$\tau_{c0}\omega_{\rm KTI,s}$ [c]
`S1K0`	0.365	1.5	3	0.12	0.06
`S1K025`	0.365	1.0	2	0.08	0.04
`S1K1`	0.365	1.0	2	0.08	0.03
`S2K0`	0.344	2.3	4	0.16	0.14
`S2K025`	0.344	2.3	4	0.16	0.13
`S2K025L`	0.344	2.3	9	0.16	0.11
`S2K1`	0.344	2.3	2	0.09[d]	0.07

[a] f_c is the ratio of the original current sheet width to the current sheet width after the readjustment phase.

[b] $N_{\rm FR}$ is the number of flux ropes per current sheet initially formed in the simulation.

[c] The two values of $\tau_{c0}\omega_{\rm KTI}$ are the normalized theoretical (subscript "t") and simulated (subscript "s") growth rates of the tearing mode.

[d] This growth rate is based on the observed mode wavelength, which is significantly longer than the fastest-growing wavelength.

is present in Harris current sheets as well as in our simulations, but is masked by the more rapid effects of kink and oblique modes.

We proceed to calculate the theoretical and actual tearing mode growth rates in each simulation. In Table 3.2 we show the two rates which we label $\omega_{\rm KTI,t}$ and $\omega_{\rm KTI,s}$, along with the initial particle drift velocity β_0, the initial current sheet compression factor f_c, and the number of flux ropes per current sheet determined by visual inspection $N_{\rm FR}$. Because the fastest-growing tearing wavelength computing from the theoretical result in Equation (3.12) is $\sim 10.8\,\lambda$ and the box size in all but simulation `S2K025L` is $20\,\lambda_0$, the number of flux ropes per current sheet produced by the tearing mode should be $N_{\rm FR} = 1.9 f_{\rm c}$; the number should be twice as large in `S2K025L`. The actual number of flux ropes formed in the simulation is agreement with this prediction, with the exception of the simulation `S2K1` with the strongest guide field. This run produced only 2 flux ropes per current sheet even though one expects 4. As discussed in Daughton et al. (2011), the strong guide field may prefer oblique rather than pure tearing modes, but the former grow the fastest at long wavelengths $\sim 33\,\lambda_{\rm p}$, about half the box size of the simulation, and the finite box size interferes with ability of arbitrary oblique modes to grow.

The values of $\omega_{\rm KTI,t}$ for the various runs are shown in Table 3.2; it should be noted that these are upper limits to the possible tearing mode growth rates, because any measured growth rates are found from estimates in the nonlinear regime. We also compute the simulated growth rates $\omega_{\rm KTI,s}$ of the tearing mode in our simulations by calculating the growth rate of the average reconnected magnetic field perturbation B_x in the current sheet. Growth rates of the KTI calculated with the latter method should be higher than those obtained by examining a single Fourier mode, as we did in Section 3.3.3, because multiple Fourier modes contribute to the amplitude of the perturbation. We find that the theoretical and simulated growth rates were relatively similar in all simulations, with theoretical rates being at most double the simulated rates. Zenitani & Hoshino (2007) identified a similar discrepancy between the theoretical and simulated growth rates. The growth rates also show that decreased guide field κ, and especially higher magnetization σ, lead to higher tearing mode growth rates both in theoretical calculations and in the simulations. This reinforces the effects of the variation in N_{FR}, which has a similar dependence on σ and κ.

We have shown that the tearing mode is responsible for producing magnetic flux ropes in most of our simulations. The tearing mode can also explain many of the differences between the simulations. The runs with higher σ and smaller κ have larger values of f_{c}, which leads to higher tearing mode growth rates and the formation of a larger number of flux ropes N_{FR}. The faster evolution of runs with higher σ and smaller κ can be explained by the faster growth of the tearing mode in such runs. Because runs with higher σ and smaller κ are also associated with a larger N_{FR}, more flux rope merging can take place in such runs. Flux rope merging is strongly associated with energy transfer and particle acceleration in our simulations, so the larger energy transfer in runs with higher σ and smaller κ can be explained by the larger N_{FR} in those runs. This suggests that a large portion of the variation between runs is a result of the effect of σ and κ on the tearing mode.

3.4.2 Particle Energization

Various particle energization channels have been identified in PIC simulations of plasma reconnection (see, e.g., Oka et al., 2010b, and references therein). Three specific loci where particle energization was detected include: inside or near the diffusion region containing reconnection X-lines, in the magnetic islands (or plasmoids, flux ropes) where the reconnected magnetic flux accumulates, and between an X-line and the edge of a flanking island, where the plasma flowing out of the X-line first encounters a strong magnetic field gradient.

A number of studies point to the conclusion that significant energization occurs near the primary (i.e., largest, or highest rank) X-lines. There, the electric field, which is perpendicular to the reversing field and is aligned with the current density vector in the middle of the current sheet, can accelerate particles oscillating within or across the current sheet (e.g., Zenitani & Hoshino, 2001, 2007, 2008; Lyubarsky & Liverts, 2008; Uzdensky et al., 2011; Bessho & Bhattacharjee, 2012; Cerutti et al., 2012a,b). A variant of this mechanism involves the trapping of particles in secondary magnetic islands appearing within the diffusion region of an X-line (Oka et al., 2010a); we will not discuss this as we do not observe secondary flux rope formation in our simulations.

Particles can also be accelerated outside of the X-lines, such as in primary magnetic islands. If an island contracts in the course of its relaxation to MHD equilibrium, the particles trapped inside it can be accelerated by a Fermi-type process (e.g., Drake et al., 2006, 2010; Kowal et al., 2011). Yet another location for particle acceleration is in the pileup region between the X-line and a flanking island, which is where the reconnected magnetic flux accumulates. There, relativistic Speiser motion can combine with curvature drift along the magnetic field gradient to create significant acceleration (e.g., Hoshino et al., 2001; Jaroschek et al., 2004; Zenitani & Hoshino, 2007; Pritchett, 2008; Huang et al., 2010; Liu et al., 2011).

These mechanisms can operate in reconnection sites where the plasma configuration, involving inflow into the diffusion region and outflow toward the flanking islands, is relatively stationary, or inside a single, autonomously evolving magnetic island. There may be other mechanisms that operate in course of spatial rearrangement and merging of magnetic islands. Converging islands can give rise to Fermi-type acceleration as particles bounce between them (e.g., Oka et al., 2010b; Tanaka et al., 2011). A dynamically active reconnection region containing many interacting islands can also allow Fermi-type stochastic particle acceleration (Drake et al., 2010; Hoshino, 2012). It is worth noting that even if particle acceleration is not directly driven by the merging, it is normally most active during island merging episodes in the course of a reconnection event (Jaroschek et al., 2004; Pritchett, 2008).

The presence of a guide field has a significant effect on particle acceleration in reconnection (Zenitani & Hoshino, 2008). Without a guide field, the direct X-line acceleration was typically found to be less efficient than the other mechanisms discussed. However, with a guide field, the effectiveness of acceleration in the pileup region (Huang et al., 2010) and inside magnetic islands (Fu et al., 2006) was diminished, thus leaving X-line acceleration of particles on Speiser orbits as the dominant mechanism. Our results are consistent with this conclusion.

To permit an interpretation of the nonthermal radiation spectra in observed high-energy astrophysical sources in terms of synchrotron and inverse-Compton radiation, a relatively hard power law particle energy spectrum, which places a significant portion of the total energy in high-energy particles, is required. In most PIC simulations including the present work, reconnection produced energized particle populations, but whether the populations represented genuine power laws tails has remained a matter of interpretation. Alternatively, an energized population can be interpreted as a composite of one or more thermal sub-populations, each at a different temperature, for example. Two dimensional simulations typically provide the dynamic range to make a tentative distinction between a thermal or a power law spectrum in an energized population. In most three-dimensional simulations, however, such a determination is dubious. In what follows, we discuss the characteristics of the particle energy spectra in the literature and provide a comparison with our results. Then, we briefly reflect on the expected nature of particle energy spectra produced by systems experiencing magnetic reconnection.

Most investigations involving PIC simulations of plasma reconnection have interpreted a section of the high energy tail of the particle energy spectrum as a power law $dN/d\ln\gamma \propto \gamma^{\alpha}$. In the X-line region of two dimensional simulations, spectra with power law indices as hard as $\alpha \approx -1$ have been reported (Zenitani & Hoshino, 2001, 2007; Jaroschek et al., 2004; Bessho & Bhattacharjee, 2007; Lyubarsky & Liverts, 2008). The spectrum in the whole simulation box also contains a power law component, but with a softer index of $\alpha \sim -2.5$. In their two- and three-dimensional simulations of shock-induced reconnection, Sironi & Spitkovsky (2011) find that as long as the region containing a reversing magnetic field that can undergo reconnection is reasonably large compared to the thickness of the reconnection layer, the particle energy spectrum at high energies is a power law with an index $\alpha = -1.5$ over a decade in energy. Other three dimensional investigations have identified relatively soft power law spectra with indices $\alpha \sim -3$ both in the X-line region and in the entire simulation box (Jaroschek et al., 2004; Zenitani & Hoshino, 2008).

Other investigations have interpreted particle energy spectra in terms of multiple thermal and other exponentially truncated populations (e.g., Oka et al., 2010b). In their two-dimensional simulation of reconnection in an initially non-relativistic plasma at temperature $T = 0.15\,m_e c^2$, Cerutti et al. (2012b) find a new thermal population at temperature $T \sim 4\,m_e c^2$. In a three-dimensional simulation beginning with a relativistic plasma at temperature $T = m_e c^2$, Liu et al. (2011) find that a thermal population with $T \approx 2.3\,m_e c^2$ is produced. In a recent work presenting two-dimensional simulations without a guide field, Bessho & Bhattacharjee (2012) detect a spectrum of the form $dN/d\gamma \propto \gamma^{-1/4} \exp(-a\gamma^{1/2})$, where a is a constant of the order of unity.

Among the cited descriptions of particle populations energized by reconnection, the spectra observed in our simulations bear resemblance with those involving multiple thermal sub-populations. For example, the spectrum at the end of simulation `S2K025L` can be modeled with three thermal components at temperatures similar to those found in Liu et al. (2011); specifically, the two energized sub-populations have temperatures $T = 2.1\,m_e c^2$ and $T = 3.5\,m_e c^2$. Neither a power law nor the Bessho & Bhattacharjee (2012) spectral form present a good fit to the spectrum in this simulation. It is important to note that the continued evolution of the particle energy spectrum we observe after the current sheet has become disordered, and the likely additional evolution that would be taking place in an even larger simulation, mean that the spectral form has not converged. For example, it is possible that in addition to the two thermal components that we have detected, additional such populations at still higher temperatures would appear in larger simulations, and that the combination of those would constitute a hard spectral tail, although it is uncertain that this tail would take the form of a power law.

In PIC simulations, distinguishing between the various forms of energized particle spectra is difficult due to dynamic range limitations. Another complication is how the spectrum of the non-energized, yet possibly adiabatically heated background plasma is to be subtracted to isolate the genuine nonthermal component. Insight can separately be gained by examining test particle trajectories in magnetic field geometries modeling reconnection layers. Particle ballistics in X-line as well as magnetic island geometries introduced at various levels of approximation has been investigated in numerous studies (e.g., Zenitani & Hoshino, 2001; Larrabee et al., 2003; Bessho & Bhattacharjee, 2012; Cerutti et al., 2012a). Often, the models entail transport terms describing particle acceleration while confined in a single X-line region or an island, and other terms describing particle escape and the termination of acceleration. Then, the terminal energy spectrum of the escaping particles is obtained by taking the input energy spectrum and deterministically transforming it by the transport terms. The crudest such models have suggested that the terminal spectrum could be a power law, e.g., of the form $dN/d\gamma \propto \gamma^{\alpha}$ with $\alpha = -(2/\pi)B_{\mathrm{rec}}/E_{\mathrm{RR}} \sim -(2/\pi)c/v_{\mathrm{A}}$, where B_{rec} is the reconnected magnetic field and E_{RR} is the electric field in the reconnection region discussed in Section 3.3.5 (Zenitani & Hoshino, 2001). Models taking a more detailed accounting of the transport of particle phase space coordinates and the kinematics of escape, however, can instead imply distinctly non-power-law, softer spectra (e.g., Bessho & Bhattacharjee, 2012). It is also important to note that the breaking of translational invariance in the direction of the current, e.g., by three dimensional phase decoherence (see Section 3.3.1), may limit particle acceleration and prevent the production of the same energized particle tail produced in two dimensional geometries.

Power-law spectra are generically expected in acceleration processes entailing stochasticity. A prime example is the linear diffusive shock acceleration (DSA) in which the random outcome of particle scattering in the shock downstream determines whether the particle will return to the shock upstream and be subjected to another acceleration cycle (Bell, 1978; Blandford & Ostriker, 1978). In contrast, the acceleration in an idealized time-independent and two-dimensional X-line region that lacks substructure in the form of secondary and embedded islands, should be deterministic. Therefore the models of Zenitani & Hoshino (2001), Larrabee et al. (2003), and Bessho & Bhattacharjee (2012), can be thought of as producing a power-law-like spectrum only "by coincidence." Realistic reconnection regions should plausibly allow particles to be accelerated in stochastic fashion as they contain time dependence and structure on multiple scales; they may also have a mechanism for returning particles that have escaped an acceleration site into another such site (see, e.g., Drake et al., 2006, 2010; Kowal et al., 2011; Hoshino, 2012). Since in this work we have found that the reconnection layer transitions into a disordered network of interacting flux ropes, it will be particularly interesting to investigate, in subsequent study, if this network allows for stochastic acceleration of particles in intermittent, secondary reconnection sites that may appear in the course of the flux rope network evolution.

3.5 Conclusions

In this paper, we carried out three dimensional PIC simulations of magnetic reconnection in a relativistic pair plasma with varying guide field strength. Plasma magnetizations, expressed in terms of the magnetic to kinetic pressure ratio, were of the order of unity. The initial conditions differed from the usual Harris sheet configuration by not having a large density contrast between the center of the current sheet and the background plasma. We investigated the growth of unstable kinetic modes in the current sheet, as well as the nonlinear development of a three dimensional flux rope network. We also investigated the character and efficiency of particle energization. Our main results can be summarized as follows:

The current sheets in all simulations develop significant magnetic reconnection accompanied with conversion of magnetic to particle kinetic energy. With the aid of Fourier decomposition, we ascertained that in all runs but Run `S1K0`, the linear tearing mode is dominant in the early evolution of the current sheet, and no significant growth occurs in the linear kink and oblique modes. The nonlinear development of the tearing mode produces a chain of flux ropes separated by primary X-lines. The flux ropes merge in hierarchical fashion whereby the merging time scale is proportional to the flux rope separation. During this phase magnetic reconnection takes place at the X-lines. We find that the dimensionless reconnection rates $\sim (0.05-0.08)$ and the maximum outflow speeds $\approx 0.4c \sim v_{\mathrm{A}}/2$ in our simulations are similar to those detected in other three dimensional simulations of reconnection in pair plasmas. We also find that spatial variation of an off-diagonal component of the pressure tensor is responsible for the breaking of flux freezing at the X-lines, consistent with existing results.

While the hierarchical flux rope merging process initially appears similar to that found in two dimensional simulations, in fact it is three-dimensional from the outset. This is because a lack of initial phase coherence in the linear tearing mode on scales larger than those allowed by causality breaks translational invariance in the direction of the initial current flow. The flux ropes form a topologically interconnected, dynamically evolving network. Dynamical interaction between neighboring flux ropes is provided by magnetic tension forces. With time, the flux ropes break up into segments with more isotropic orientations. The strongly three-dimensional character of the reconnection layer seems to suggest that global reconnection models invoking quasi-two-dimensional plasmoid hierarchies (e.g., Shibata & Tanuma, 2001; Fermo et al., 2010; Uzdensky et al., 2010) require revision to account for the inter-plasmoid magnetic linkage and isotropization of plasmoid orientations.

The larger flux ropes produced during flux rope merging contain substructure down to plasma skin depth scales which is reflected in embedded, twisted and braided current filaments and sheets. Overall, this substructure is force-free and evolves relatively slowly. However, isolated sites within the evolved flux rope network contain spatially and temporally intermittent sites characterized by strong nonideal conditions $\mathbf{E} \cdot \mathbf{B} \neq 0$ where a change of magnetic connectivity continues to take place even after flux rope merging has saturated on length scales equal to the size of the computational box. This intermittency may produce the observed variability of nonthermal emission in systems in which the emitting particles are energized by magnetic reconnection.

During the early, ordered flux rope merging phase, particles are accelerated to high Lorentz factors by the electric field in primary X-lines; the trajectories of these particles are well described by Speiser orbits. Particles continue to be energized in the later, disordered phase we identify in our largest simulation, but we leave the analysis of energization in the disordered regime to a subsequent investigation.

Simulations with higher magnetization and lower guide field strength exhibit greater and faster energy conversion and particle energization. The efficiency of particle energization measured in terms of the energy in the accelerated particles per unit magnetic energy dissipated in the simulation is an increasing function of the guide field strength for $\sigma = 1$, which can be interpreted as resulting from a decreasing plasma compressibility with increasing guide field. The final particle energy spectrum in the largest simulation is best fit by the inclusion of new thermal components at temperatures $2.1\, m_e c^2$, and $3.5\, m_e c^2$, in addition to the initial thermal component with temperature $m_e c^2$. We, however, acknowledge that a larger size or longer duration simulation is likely to produce a still more pronounced energized component, possibly even a population described with a power law spectrum.

Energetic positrons (electrons) with Lorentz factors $\gamma > 30$ are moderately beamed in (opposite to) the direction of the initial current flow with median inclinations of $\sim 30^\circ - 40^\circ$. The degree of beaming is determined by a particle's energy gain during acceleration. We speculate that more highly magnetized plasmas and reconnection sites with larger size X-line regions should give rise to stronger beaming.

In this work, we have investigated a narrow range of magnetizations with $\sigma \sim \mathcal{O}(1)$, but astrophysical reconnection sites can also have high magnetizations $\sigma \gg 1$. We can speculate about the applicability of our results in the latter limit. The linear tearing mode responsible for the initiation of reconnection is insensitive to the degree of magnetization far from the current sheet. The phase decoherence that produces the initial breakdown of translational invariance is determined by the tearing mode growth time and should thus also persist at high magnetizations. Therefore, we expect the qualitative structure of the reconnection region at higher values of σ to be similar to that found in our simulations.

The primary effect of high magnetization is that the Alfvén velocity approaches the speed of light, which could give rise to ultrarelativistic outflows from the X-line region. In such outflows the inertial term of the generalized Ohm's law becomes important in the breaking of flux freezing (Hesse & Zenitani, 2007). This in turn may increase the dimensionless reconnection rate $r_{\rm rec}$ relative to the value found in our simulations. It remains to be seen whether the associated reconnection process is more or less intermittent. An increased magnetization is likely to increase the efficiency and the degree of beaming in particle energization.

We consider these results and the immediate questions they raise an incremental step in the development of a multiscale view of collisionless plasma self-organization during magnetic reconnection. Further work is clearly required to place our key finding, the evolution of the simulated, periodic reconnection layer into a disordered network of interacting magnetic flux ropes, in the macroscopic context of a realistic reconnection site characterized by outflow boundary conditions and altogether different field line asymptotics at large distances from the X-line. It will be particularly interesting to see if the reconnected-flux-carrying outflow from the macroscopic X-line will possess the disordered, interlinked magnetic field topology we observe and investigate what will be the character of magnetic fluctuations in the outflow.

Chapter 4

The Role of the Magnetorotational Instability in the Sun[1]

One of the major unsolved problems of stellar evolution is the effect of rotation, especially differential rotation, on the magnetic field structure of stars, and the feedback of that magnetic field on the stellar structure and evolution (Maeder, 2009, and references therein). While the detailed interior magnetic field structure is not known for any star, Global Oscillations Network Group (GONG, Howe et al., 2005) and Michelson Doppler Imager (MDI, Christensen-Dalsgaard et al., 1996) helioseismology has allowed the detailed calculation of the rotation profile of the Sun. In the radiative zone at $r < 0.69r_\odot$, the Sun has approximately solid-body rotation. In the convection zone at $r > 0.72r_\odot$ the rotation rate can be primarily described as a slowly increasing function of spherical θ, except near the solar surface at $r > 0.9r_\odot$ where a strong radial shear layer is present with $\partial \ln \Omega / \partial \ln r < 0$. The radius of the transition from radiative to convective transport is $r \approx 0.713r_\odot$, and the transition is associated with a very strong radial shear layer with width $\sim 0.02r_\odot$ known as the tachocline, which has $\partial \ln \Omega / \partial \ln r < 0$ close to the poles and $\partial \ln \Omega \partial \ln r > 0$ close to the equator. The tachocline's central radius is $r \sim 0.70r_\odot$, and varies varies slightly

---

[1] A large part of the material in this chapter is taken from the manuscript Kagan, D., Wheeler, J. C., 2013, which has been submitted to the Astrophysical Journal.

with latitude; it is likely located entirely in the radiative region near the equator, but it may reach significantly into the convective region closer to the poles (Basu & Antia, 2001). Combining this observed rotation profile with observed solar surface magnetic fields promises significant progress in understanding the origin of interior solar magnetic fields.

Observations of surface solar magnetic fields indicate the presence of fields on large spatial scales that vary with the solar cycle and are associated with active regions and eruptive events. The most popular approach to understanding the origin of these large-scale fields involves α–Ω dynamo models associated with the tachocline (e.g., Ossendrijver, 2003, and references therein). In these models, nonaxisymmetric instability or turbulence produces relatively small-scale, nonaxisymmetric poloidal fields from toroidal fields in the lower convection zone and convective overshoot region of the upper tachocline (the α effect). Field line wrapping by strong differential rotation (the Ω effect) in the tachocline then stretches poloidal fields into large-scale toroidal fields and completes the dynamo loop. The process behind the Ω effect which produces toroidal fields by stretching poloidal fields is well understood, but there are many candidate mechanisms for the α effect that produces poloidal fields from toroidal fields. One possible way of producing an α effect is a kinematic mean-field dynamo based on the Parker (1955) mechanism, in which small-scale, convective, nonaxisymmetric turbulence causes toroidal fields to be passively advected into poloidal fields. Passive advection is also used to produce poloidal fields in Babcock-Leighton models, in which the source of the α effect is the twisting of large-scale toroidal field structures under the influence of the coriolis force as they rise through the convection zone. An alternative possibility for the α effect is that a local or global hydrodynamic or magnetohydrodynamic (MHD) instability in the tachocline region is directly responsible for the production of poloidal field. A global hydrodynamic instability may be present in the tachocline as a result of the latitudinal differential rotation (Dikpati & Gilman, 2001). Local MHD instabilities that can be driven by strong radial shear in the tachocline include the Tayler-Spruit instability (Pitts & Tayler, 1985; Spruit, 2002) and the magnetorotational instability (MRI), which is the subject of this paper.

In addition to the large-scale fields associated with solar activity, small-scale fields have been observed on the quiet sun that do not vary with the solar cycle. These fields are typically thought to result from small-scale kinetic dynamo action due to convective turbulence Martínez Pillet (e.g., 2013, and references therein), probably in situ in the upper convection zone (**?**). The presence of a strongly radial shear near the surface of the Sun indicates that MHD instabilities like the MRI and the Tayler-Spruit instability may also grow in the convective zone and play a role in the origin of these magnetic fields.

4.0.1 MRI

The magneto-rotational instability (MRI; Velikhov, 1950; Chandrasekhar, 1960; Acheson, 1978; Balbus & Hawley, 1991, 1998), which is driven by a negative radial angular velocity gradient, has been thoroughly explored in the context of accretion disks, but it also applies to quasi-spherical objects, e.g. stars (Balbus & Hawley, 1994). A general dispersion relation and associated instability criteria encompassing the MRI and other instabilities for the nonaxisymmetric, diffusive case with finite resistivity, viscosity, and conductivity, is presented in Acheson (1978). That is the relation that should be applied in stars, but it remains cumbersome to employ and relatively unexplored. Therefore, other dispersion relations have been derived that include small-scale magnetohydrodynamic (MHD) shear modes that have similar properties to the MRI modes in accretion disks, as well as other large-scale hydrodynamic modes associated with convection and shear. Balbus & Hawley (1994) (see also Chanmugam, 1979) derived a dispersion relation including the MRI for axisymmetric, non–diffusive conditions with rotation restricted to cylinders. Balbus (1995) generalized the dispersion relation to a general distribution of angular velocity, $\Omega(\varpi, Z)$, where ϖ is the cylindrical radius and Z is the vertical coordinate, while Menou et al. (2004) derived a dispersion relation including the effects of viscous, thermal, and magnetic diffusivities. Kim & Ostriker (2000) and Masada et al. (2006) considered the nonaxisymmetric, non–diffusive modes including the MRI in the context of shearing winds and proto-neutron stars, respectively. Masada et al. (2007) explored a nonaxisymmetric dispersion relation containing the MRI including the effects of the three classical diffusivities and neutrino diffusion in the context of proto-neutron stars with spherically symmetric rotation profiles.

4.0.2 MRI in the Sun

The fastest-growing modes of the MRI typically have $kv_{\rm A} \sim \Omega$, where $v_{\rm A} \equiv B/\sqrt{4\pi\rho}$ is the Alfvén speed, k is the poloidal wavenumber, B is the magnetic field and ρ is the density. The length scale of these modes is therefore $\lambda \sim 2\pi v_A/\Omega$. For a seed field $B \sim 1G$ in the tachocline, where roughly $\rho \sim 0.1$g cm^{-3} and $\Omega \sim 10^{-6}$rad s^{-1}, we obtain $\lambda \sim 10^5$cm $\sim 10^{-5} r_\odot$. Because the fastest-growing modes of the MRI are so small relative to the solar radius, it is currently impracticable to resolve the MRI in global simulations of the Sun. As a result, most analysis of the MRI in stars has been confined to application of local MRI dispersion relations. The action in the Sun of the triply-diffusive axisymmetric instability derived by Menou et al. (2004) has been investigated by Menou et al. (2004) in the radiative region, by Parfrey & Menou (2007) in the solar tachocline, and by Masada (2011) in the tachocline and convective zone in isolation from the effects of convection. Menou et al. (2004) note that in their extensive stability analysis in both the inviscid and perfect-conductor double-diffusive limits *any* level of negative differential rotation is destabilized by a combination of diffusion–free (along spherical shells) and double–diffusive (across spherical shells) modes. They found, however, that even a relatively small viscosity could add some stability in the triply–diffusive case for differential rotation between shells, $d\ln\Omega/d\ln\varpi < 0$, so that this case, which is the most important one in the Sun, must be considered quantitatively.

Parfrey & Menou (2007) investigated the growth rate of modes in the stably stratified tachocline, where only MRI modes exist. Using a semianalytical prescription for the differential rotation along the tachocline, they found that the regions of $\theta \lesssim 60^{\circ}$ are formally unstable to the MRI, but significant growth of instability occurs only for $\theta < 53^{\circ}$. Parfrey & Menou (2007) concluded that the turbulence associated with the MRI at high latitudes disrupts the formation of large scale magnetic fields. They argued that such large scale magnetic fields can only form at lower latitudes by more traditional solar dynamos operating in the tachocline.

Masada (2011) employed the triply–diffusive dispersion relation for the axisymmetric instability of Menou et al. (2004) combined with rotation profiles determined from helioseismology and a standard model of the Sun to calculate the growth rate of the MRI throughout the Sun. He neglected the destabilizing effects of thermal buoyancy, but included the effects of stable stratification, excluding all non-MRI modes from the analysis. He found that unstable modes existed in the tachocline at the same latitudes found by Parfrey & Menou (2007) as well as near the solar surface at low latitudes $\theta > 45^{\text{o}}$. He argued that the calculated growth of instability near the solar surface was unrealistic, because the strong convection present there would disrupt this growth.

In this paper, we consider the full triply–diffusive, nonaxisymmetric magneto-convective instability in a model of the current Sun; we thus assume that both thermal buoyancy and shear effects contribute to the instability. One may question whether it makes sense to do a linear stability analysis in a solar model, since turbulent motions in the convective region imply that the stationary background necessary to derive a dispersion relation is not present (Masada, 2011); however, recent helioseismic results indicate that large-scale turbulent velocities in the upper convection region of the Sun are typically quite small, of order 0.01 km s^{-1} (Hanasoge et al., 2010, 2012). This suggests that application of a linear analysis is possible even near the solar surface. With these assumptions we find that there are indeed parts of the convective region of the Sun where the growth rate of MRI modes is more rapid than the growth of convective modes due to thermal buoyancy.

The paper is organized as follows. In section 4.1, we derive a dispersion relation for the triply diffusive nonaxisymmetric instability including the MRI and discuss the relevant instability criteria for that dispersion relation in the Sun. Section 4.2 describes our methodology for calculating the growth rate of modes in the Sun. Section 4.3 presents our results. Section 4.4 compares our findings to previous work on the MRI in the Sun, and discusses the nonlinear saturation of shear modes including the MRI and convection in the Sun. Finally, Section 4.5 reviews our main conclusions.

4.1 The Nonaxisymmetric MRI

4.1.1 Dispersion Relation

We now calculate the growth rates of nonaxisymmetric, diffusive modes including the MRI that may be present in the Sun. To do this, we carry out a Wentzel—Kramers—Brillouin (WKB) perturbation analysis of the equations of magnetohydrodynamics (MHD) under the assumption that pressure perturbations are negligible except in the momentum equation where they are coupled to buoyancy effects (i.e., the Boussinesq approximation). We neglect composition gradients, because even the outer portions of the radiative region in the Sun are expected to have a homogeneous composition, and mixing in the convective region guarantees this homogeneity. The MHD equations under these approximations are

$$\nabla \cdot \mathbf{v} = 0, \tag{4.1}$$

$$\left(\frac{\partial}{\partial t}-\nu\nabla^2\right)\mathbf{v}+(\mathbf{v}\cdot\nabla)\mathbf{v}=-\frac{1}{\rho}\nabla\left(P+\frac{\mathbf{B}^2}{8\pi}\right)$$

$$+\frac{1}{4\pi\rho}(\mathbf{B}\cdot\nabla)\mathbf{B}+\mathbf{g}, \tag{4.2}$$

$$\left(\frac{\partial}{\partial t}-\eta\nabla^2\right)\mathbf{B}=(\mathbf{B}\cdot\nabla)\mathbf{v}-(\mathbf{v}\cdot\nabla)\mathbf{B}. \tag{4.3}$$

$$\left(\frac{\partial}{\partial t}+\mathbf{v}\cdot\nabla\right)\ln\frac{P}{\rho^\gamma}=\xi\nabla^2\tau, \tag{4.4}$$

In these equations, $\mathbf{v}$ is the fluid velocity, ν is the kinematic viscosity, η is the magnetic resistivity, ξ is the thermal diffusivity, $\mathbf{g}$ is the acceleration due to gravity, and $\tau \equiv T/T_0$ is a dimensionless temperature parameter normalized to the local equilibrium temperature T_0. Note that in Equation (4.4), we have assumed that the fluid may be treated as an ideal gas with adiabatic index γ. To complete our equation set, we may relate $\mathbf{g}$ to known thermodynamic quantities by assuming the initial mass distribution is in hydrostatic equilibrium. Because the equilibrium gravitational force is much larger than the equilibrium magnetic and shear forces, we may express $\mathbf{g}$ in cylindrical coordinates (ϖ,ϕ,Z) as:

$$\mathbf{g}=\left(-\frac{1}{\rho}\frac{dP}{d\varpi},0,-\frac{1}{\rho}\frac{dP}{dZ}\right). \tag{4.5}$$

We now perform a local WKB analysis in cylindrical coordinates (ϖ,ϕ,Z), assuming that the perturbations are of the form $\delta \propto \exp\{i(k_\varpi\varpi+m\phi+k_Z Z-\sigma t)\}$, where m is an integer. We express oscillation frequencies in terms of $\omega\equiv\sigma-m\Omega$, which is the relevant oscillation frequency for disturbances in the rotating frame. Accounting for the effects of dissipation on these oscillations, we then introduce the variables

$$\omega_\alpha = \omega + i\alpha k^2 \;\; (\alpha = \xi, \eta, \nu). \tag{4.6}$$

In order to use this WKB form for the perturbations, we must make the local approximation $m/\varpi \ll k_\varpi, k_Z$. As a result, the perturbation in the total pressure is negligible in the ϕ component of the momentum equation, and the nonaxisymmetric component of the continuity equation is negligible. As noted above, we also apply the Boussinesq approximation by setting $\delta P = 0$ in all equations but the momentum equation, Equation (4.2). We assume the local equilibrium magnetic field is uniform for simplicity; for weak initial fields, gradients in these fields are unlikely to have strong effects. Finally, we make the assumption that the equilibrium magnetic field is primarily toroidal; i.e., $B_\phi \gg B_\varpi, B_Z$. Therefore, although we neglect B_ϖ/ϖ relative to $k_\varpi B_\varpi$ due to the local approximation, we do *not* neglect B_ϕ/ϖ and mB_ϕ/ϖ relative to $k_\varpi B_\varpi$ in calculating the linearized equations.

Keeping only linear order terms in equations (4.1)–(4.4), we find the following 8 equations for the 8 perturbed quantities (the three components of $\delta\mathbf{B}$ and $\delta\mathbf{v}$, δP, and $\delta\rho$):

$$k_\varpi \delta v_\varpi + k_Z \delta v_Z = 0, \tag{4.7}$$

$$\begin{aligned} i\omega_\nu \delta v_\varpi + 2\Omega \delta v_\phi = & ik_\varpi \left(\frac{\delta P}{\rho} + \frac{\mathbf{B}\cdot\delta\mathbf{B}}{4\pi\rho} \right) - \frac{i(\mathbf{k}\cdot\mathbf{B})}{4\pi\rho}\delta B_\varpi \\ & + \frac{2}{\varpi}\frac{B_\phi \delta B_\phi}{4\pi\rho} - \frac{\delta\rho}{\rho^2}\frac{dP}{d\varpi}, \end{aligned} \tag{4.8}$$

$$\begin{aligned} i\omega_\nu \delta v_\phi - \frac{\kappa^2}{2\Omega}\delta v_\varpi - \varpi\frac{d\Omega}{dZ}\delta v_Z =& - \frac{i(\mathbf{k}\cdot\mathbf{B})}{4\pi\rho}\delta B_\phi \\ & - \frac{1}{\varpi}\frac{B_\phi \delta B_\varpi}{4\pi\rho}, \end{aligned} \tag{4.9}$$

$$\begin{aligned} i\omega_\nu \delta v_Z =& i k_Z \left(\frac{\delta P}{\rho} + \frac{\mathbf{B}\cdot\delta\mathbf{B}}{4\pi\rho}\right) \\ & - \frac{i(\mathbf{k}\cdot\mathbf{B})}{4\pi\rho}\delta B_Z - \frac{\delta\rho}{\rho^2}\frac{dP}{dZ}, \end{aligned} \tag{4.10}$$

$$\omega_\eta \delta B_\varpi = -(\mathbf{k}\cdot\mathbf{B})\delta v_\varpi \ , \tag{4.11}$$

$$\begin{aligned} i\omega_\eta \delta B_\phi + \frac{B_\phi}{\varpi}\delta v_\varpi =& - \frac{d\Omega}{d\ln\varpi}\delta v_\varpi \\ & - \varpi\frac{d\Omega}{dz}\delta v_Z - i(\mathbf{k}\cdot\mathbf{B})\delta v_\phi, \end{aligned} \tag{4.12}$$

$$\omega_\eta \delta B_Z = -(\mathbf{k}\cdot\mathbf{B})\delta v_Z, \tag{4.13}$$

$$i\omega_\xi \gamma \frac{\delta\rho}{\rho} + \delta v_\varpi \frac{d\ln P\rho^{-\gamma}}{d\varpi} + \delta v_Z \frac{d\ln P\rho^{-\gamma}}{dZ} = 0, \tag{4.14}$$

where we have defined the epicyclic frequency to be:

$$\kappa^2 \equiv \frac{1}{\varpi^3}\frac{d\Omega^2\varpi^4}{d\varpi} = 4\Omega^2 + \frac{d\Omega^2}{d\ln\varpi} \tag{4.15}$$

Note that in Equation (4.14), we have eliminated the temperature perturbation using the relation $\delta\rho/\rho = -\delta T/T$, which may be derived by combining the Boussinesq approximation $\delta P = 0$ with the ideal gas law $P = \rho k_B T$.

Combining equations (4.7)-(4.14), we find the triply-diffusive nonaxisymmetric dispersion relation including the MRI:

$$\begin{aligned}\frac{k_{\rm pol}^2}{k_Z^2}\widetilde{\omega}_{\eta\nu}^4 - \widetilde{N}^2\frac{\omega_\eta}{\omega_\xi}\widetilde{\omega}_{\eta\nu}^2 - \widetilde{\kappa}^2\widetilde{\omega}_\eta^2 - 2{\omega_{{\rm A}\phi}}^2\widetilde{\omega}_{\eta\nu}^2 & \\ -4(\mathbf{k}\cdot\mathbf{v}_{\rm A})^2\Omega^2\left(1+\left[\frac{\omega_{{\rm A}\phi}}{\Omega}\right]^2\right) & \\ -\frac{(\mathbf{k}\cdot\mathbf{v}_{\rm A})\omega_{{\rm A}\phi}}{\Omega}(\omega_\eta[4\Omega^2+\widetilde{\kappa}^2]+\omega_\nu[4\Omega^2-\widetilde{\kappa}^2]) = 0\ , & \qquad (4.16)\end{aligned}$$

where

$$k_{\rm pol}^2 = k_\varpi^2 + k_Z^2, \qquad (4.17)$$

$$\widetilde{\omega}_{\eta\nu}^2 \equiv \omega_\eta\omega_\nu - (\mathbf{k}\cdot\mathbf{v}_{\rm A})^2\ , \qquad (4.18)$$

$$\widetilde{\omega}_\eta^2 \equiv \omega_\eta^2 - (\mathbf{k}\cdot\mathbf{v}_{\rm A})^2\ , \qquad (4.19)$$

$$\omega_{{\rm A}\phi} \equiv \frac{(\mathbf{k}\cdot\mathbf{v}_{\rm A})_\phi}{m} = \frac{B_\phi}{\varpi\sqrt{4\pi\rho}}, \qquad (4.20)$$

$$\widetilde{N}^2 \equiv -\frac{1}{\rho\gamma}(\mathcal{D}P)\,\mathcal{D}\ln(P\rho^{-\gamma}), \qquad (4.21)$$

$$\widetilde{\kappa}^2 \equiv -\frac{1}{\varpi^3}\,\mathcal{D}(\varpi^4\Omega^2) = \kappa^2 - \frac{k_\varpi}{k_Z}\varpi\frac{d\Omega^2}{dZ}, \qquad (4.22)$$

$$\mathcal{D} \equiv \left(\frac{k_\varpi}{k_Z}\frac{d}{dZ} - \frac{d}{d\varpi}\right). \qquad (4.23)$$

The dispersion relation (4.16) is very similar to that found by Masada et al. (2007, their Equation 30); the equations differ only because they neglect the shear in the *Z* direction, which is important in the Sun, while we neglect neutrino radiation, which has a negligible effect in the Sun. If we neglect nonaxisymmetric effects completely by setting $\omega_{\mathrm{A}\phi} = 0$, we recover the dispersion relation of Menou et al. (2004, their Equation 13).

For modes on very large length scales, which correspond to small k, this dispersion relation can be simplified further. For large-scale modes, the characteristic dissipative frequencies $k^2\xi$, $k^2\eta$, and $k^2\nu$ are small compared to the rotation rate Ω and the magnitude of the buoyancy frequency $|N|$. Therefore, for all fast-growing modes with $|\omega| \sim \Omega$ or $|\omega| \sim |N|$, from Equation 4.6 we have $\omega_\xi \sim \omega_\eta \sim \omega_\nu \sim \omega$ and from Equations (4.18) and (4.19) we have $\widetilde{\omega}^2_{\eta\nu} \sim \widetilde{\omega}^2_\eta \sim \widetilde{\omega}^2$, where we define $\widetilde{\omega}^2 \equiv \omega^2 - (\mathbf{k}\cdot\mathbf{v}_\mathrm{A})^2$. The resulting dispersion relation is then

$$\frac{k^2_{\mathrm{pol}}}{k^2_Z}\widetilde{\omega}^4 - (\widetilde{N}^2 + \widetilde{\kappa}^2 + 2\omega_{\mathrm{A}\phi}{}^2)\widetilde{\omega}^2$$
$$-4(\mathbf{k}\cdot\mathbf{v}_\mathrm{A})^2\Omega^2\left(1 + 2\frac{\omega_{\mathrm{A}\phi}\omega}{(\mathbf{k}\cdot\mathbf{v}_\mathrm{A})\Omega} + \left[\frac{\omega_{\mathrm{A}\phi}}{\Omega}\right]^2\right) = 0. \qquad (4.24)$$

For weak initial fields such that $\omega_{\mathrm{A}\phi} \ll \Omega$ and $\omega_{\mathrm{A}\phi} \ll |N|$, small k also implies that the characteristic magnetic frequency $\mathbf{k}\cdot\mathbf{v}_\mathrm{A} \ll \Omega$ and $\mathbf{k}\cdot\mathbf{v}_\mathrm{A} \ll |N|$. Therefore, for fast-growing modes with $|\omega| \sim \Omega$ or $|\omega| \sim |N|$, all terms that involve $\mathbf{k}\cdot\mathbf{v}_\mathrm{A}$ and $\omega_{\mathrm{A}\phi}$ are negligible, and the dispersion relation becomes

$$\frac{k^2_{\mathrm{pol}}}{k^2_Z}\omega^2 - \widetilde{N}^2 - \widetilde{\kappa}^2 = 0\,. \qquad (4.25)$$

This final dispersion relation implies that for weak magnetic fields, large-scale modes are both adiabatic and hydrodynamic, since there is no coupling to the magnetic field or dissipation.

In stars like the Sun, the thermodynamic variables density ρ and pressure P, as well as the buoyancy frequency N, are typically functions only of spherical radius.

In these cases, $\widetilde{N}$ from may be expressed in the form

$$\widetilde{N}^2 = \left(\frac{k_\varpi}{k_Z}\cos\theta - \sin\theta\right)^2 N^2, \tag{4.26}$$

where

$$N^2 = -\frac{1}{\rho\gamma}\frac{dP}{dr}\frac{d\ln P\rho^{-\gamma}}{dr}. \tag{4.27}$$

is the square of the buoyancy frequency N, which is a function only of spherical radius.

We may also express $\widetilde{\kappa}$ in a simpler form as:

$$\widetilde{\kappa}^2 = \kappa^2 - \frac{k_\varpi}{k_Z}\varpi\frac{d\Omega^2}{dZ}. \tag{4.28}$$

In the equatorial plane of such a star, $\widetilde{\kappa} = \kappa$ and $\widetilde{N} = N$.

4.1.2 Instability Criteria

We now calculate the instability criteria for this dispersion relation that are of importance in the Sun. Because we take the initial magnetic field to be weak, we can make the approximation $\Omega \gg \omega_{\mathrm{A}\phi}$; we also eliminate the "kink-type" modes discussed by Masada et al. (2006) by focusing on modes in which $\mathbf{k}\cdot\mathbf{v}_{\mathrm{A}} \gg \omega_{\mathrm{A}\phi}$. Under this approximation, the manifestly nonaxisymmetric terms disappear, and the dispersion relation becomes

$$\frac{k_{\mathrm{pol}}^2}{k_Z^2}\widetilde{\omega}_{\eta\nu}^4 - \widetilde{N}^2\frac{\omega_\eta}{\omega_\xi}\widetilde{\omega}_{\eta\nu}^2 - \widetilde{\kappa}^2\widetilde{\omega}_\eta^2 - 4(\mathbf{k}\cdot\mathbf{v}_{\mathrm{A}})^2\Omega^2 = 0. \tag{4.29}$$

It is important to note that because a toroidal field is present, $\mathbf{k}\cdot\mathbf{v}_{\mathrm{A}}$, and therefore the dispersion relation, still has a dependence on the nonaxisymmetric wavenumber m. This dispersion relation, Equation (4.29), is identical in form to that of Menou et al. (2004) for the axisymmetric instability. In what follows, we will make use of their results in the case of spherically symmetric contours of density ρ and pressure P to calculate the instability criteria for the dispersion relation in various limits that are relevant in the Sun. Because the ordering of the diffusion parameters in the Sun is $\xi \gg \eta \gg \nu$ (see Section 4.2), the appropriate conditions for stability are those for the limit $\nu \to 0$, given by Menou et al. (2004) Equations 21, 37, 50, 56, and 62. Written in our notation, these conditions are

$$\widetilde{N}^2 + \widetilde{\kappa}^2 > 0, \tag{4.30}$$

$$\widetilde{N}^2 + \widetilde{\kappa}^2 - 4\Omega^2 > 0, \tag{4.31}$$

$$2\frac{\eta}{\xi}\widetilde{N}^2 + (1 + \frac{\eta}{\xi})\widetilde{\kappa}^2 > 0, \tag{4.32}$$

$$\frac{\eta}{\xi}\widetilde{N}^2 + \widetilde{\kappa}^2 - 4\Omega^2 > 0, \tag{4.33}$$

$$\widetilde{N}^2 > 0. \tag{4.34}$$

We will now discuss which of these stability conditions are violated in various locations in the Sun. Equation (4.26) implies that $\widetilde{N}^2$ has the same sign as N^2; therefore, there will be major differences between the stability characteristics of the dispersion relation in stably stratified regions and in convectively unstable regions, and we will treat them separately.

4.1.2.1 Stably Stratified Regions

In the radiative zone and the lower tachocline, the Sun is strongly stratified, with $N^2 \gg \Omega^2 > 0$. Equation (4.26) then implies that $\widetilde{N}^2$ is a positive definite quantity in these regions. Then, there are two possible ways in which the conditions 4.30–4.34 may be violated. Firstly, there are a small set of modes for which the wavevector is very close to being in the $\pm\theta$ direction; this corresponds to $|k_\varpi/k_Z + \tan\theta| \lesssim \Omega/N$. These modes are typically unimportant in the Sun.

The other case corresponds to $\widetilde{N}^2 \gg \Omega^2 \sim \widetilde{\kappa}^2$. In this case, we may neglect factors of η/ξ that are not multiplied by $\widetilde{N}^2$. Comparing the five stability conditions then indicates the necessary and sufficient stability criterion is

$$\frac{\eta}{\xi}\widetilde{N}^2 + \widetilde{\kappa}^2 - 4\Omega^2 > 0, \tag{4.35}$$

Following Menou et al. (2004), this criterion implies that unstable modes exist if

$$\frac{\eta}{\xi}N^2 + \frac{d\Omega^2}{d\ln\varpi} < 0, \tag{4.36}$$

or if

$$\left(\varpi\frac{d\Omega^2}{dZ}\right)^2 - 8\frac{\eta}{\xi}N^2\sin(\theta)\cos(\theta)\frac{d\Omega^2}{d\theta} > 0. \tag{4.37}$$

Note that we again neglect factors of η/ξ that are not multiplied by $\widetilde{N}^2$. The first condition represents the destabilizing influence of cylindrically radial shear, which is opposed by stable stratification; while the second represents the destabilizing influence of shear in the Z and θ directions. Both of these conditions for instability correspond to small-scale magnetohydrodynamic modes driven by shear, which can properly be called MRI modes.

4.1.2.2 Convectively Unstable Locations

In convectively unstable regions with $N^2 < 0$, Equation (4.26) implies that $\widetilde{N}^2$ is also negative. Therefore, buoyant effects always contribute to instability, and modes can be driven by the combined effects of convection and shear. The instability criteria for these modes are generally quite complicated; any or all of the conditions 4.30–4.34 may be violated. However, two limiting cases exist in which the instability criteria are more tractable. In the limiting case where $\eta/\xi \left|N^2\right| \gg \Omega^2$, rotational effects are negligible. Then, all of the stability conditions reduce to

$$\widetilde{N}^2 > 0. \tag{4.38}$$

The resulting conditions where instability can occur are

$$N^2 < 0, \tag{4.39}$$

and

$$\frac{d\Omega^2}{d\theta} > 0. \tag{4.40}$$

Because the first inequality is always satisfied in the convective zone, it is generally the important one for this case; it corresponds directly to hydrodynamic convective modes on large scales.

The other limiting case occurs when convective effects are negligible, which corresponds to $\Omega^2 \gg \left|N^2\right|$. Then, the sufficient instability criteria take the form:

$$\widetilde{\kappa}^2 - 4\Omega^2 > 0. \tag{4.41}$$

The conditions under which instability can occur are given by

$$\frac{d\Omega^2}{d\ln\varpi} < 0, \tag{4.42}$$

and

$$\frac{d\Omega^2}{dZ} \neq 0. \tag{4.43}$$

The first inequality corresponds to small-scale MRI modes that are similar to those to those found in the stably stratified regions. The second inequality is always violated unless rotation is constant on cylinders; it corresponds to large-scale hydrodynamic shear modes.

4.2 Methodology

We now calculate the growth rates of unstable modes of the triply-diffusive nonaxisymmetric dispersion relation throughout the Sun. To calculate thermodynamic variables, we compute a 1D model of the Sun using MESA (Paxton et al., 2011); we thus assume that all thermodynamic variables, such as P and ρ, are functions only of spherical radius r. We note that the one dimensional non-rotating solar model we have computed is not self-consistent with the known rotational profile of the Sun or the initially assumed magnetic fields; however, adoption of this Solar model is a necessary first step to a deeper understanding of the rotating, magnetic evolution of the Sun and other stars.

Following Menou et al. (2004), we calculate the values of the three diffusivities in the Sun from the thermodynamic quantities in the MESA model. The thermal diffusivity is dominated by radiative transport, and is given by

$$\xi = \frac{\gamma - 1}{\gamma}\frac{T}{P}\frac{16T^3}{3\kappa\rho}, \tag{4.44}$$

where κ is the radiative opacity.

The resistivity η is given by

$$\eta \approx 5.2 \times 10^{11} \frac{\ln\Lambda}{T^{3/2}}\,\mathrm{cm^2\,s^{-1}}, \tag{4.45}$$

where $\ln\Lambda$ is the Coulomb logarithm. This logarithm is given in the NRL plasma formulary as

$$\ln\Lambda \approx \begin{cases} -17.4 + 1.5\ln T - 0.5\ln\rho & T < 1.1\times 10^5\,\mathrm{K}, \\ -12.7 + \ln T - 0.5\ln\rho & T > 1.1\times 10^5\,\mathrm{K}. \end{cases} \tag{4.46}$$

after translating into cgs units.

The viscosity ν is dominated by thermal viscosity, and is given by Spitzer (2006) as

$$\nu \approx 5.2 \times 10^{-15} \frac{T^{5/2}}{\rho\ln\Lambda}\,\mathrm{cm^2\,s^{-1}}. \tag{4.47}$$

We calculate the gradients of Ω, in the convective and radiative zones using GONG helioseismic data (Howe, 2009). However, this data lacks fine spatial resolution in the tachocline, and therefore greatly underestimates the radial shear there. Therefore, in the region of the tachocline, we instead estimate the radial shear using the approximate equation (Parfrey & Menou, 2007)

$$\frac{\partial\Omega}{\partial r} = \frac{\delta\Omega_{\rm eq}}{\Delta}\left(1 - 3.56\cos^2\theta - \cos^4\theta\right), \tag{4.48}$$

where $\delta\Omega_{\rm eq} = 1.08\times10^{-7}$ rad/s is the change in angular velocity across the tachocline at the equator, and $\Delta = 0.02r_\odot$ is the width of the tachocline.

We then solve Equation (4.16) for ω as a function of location in the Sun, the wavenumber $\mathbf{k} = (k_\varpi, m/\varpi, k_Z)$, and the initial magnetic field $\mathbf{B} = (B_\varpi, B_\phi, B_Z)$. The components of the magnetic field and the wavenumber appear in the dispersion relation (4.16) solely through their contributions to k_ϖ/k_z,k^2, $(\mathbf{k}\cdot\mathbf{v}_{\rm A})^2$, $\omega_{\rm A\phi}^2$, and $(\mathbf{k}\cdot\mathbf{v}_{\rm A})\omega_{\rm A\phi}$. For values of m that obey the local approximation, the magnitude of the wavevector is $k\approx k_{\rm pol}$. Therefore, we can specify values of complete parameter space by setting k, $B_{\rm pol}$, k_ϖ/k_z, B_ϖ/B_z, m, and $R_{\rm TP}\equiv B_\phi/\sqrt{B_\varpi^2+B_Z^2}$, which is the ratio of the toroidal and poloidal magnetic field components.

Then we specify the ranges of each of these parameters; we first choose the appropriate ranges for the wavenumber k. The local approximation creates a constraint requiring that the mode wavelength $\lambda = 2\pi/k$ fits within a single local pressure scale height $H_{\rm P}$; a lower limit on λ is provided by the effects of magnetic resistivity and viscosity, which stabilize modes on scales for which $k^2\eta$, $k^2\nu \gg kv_{\rm A}$. The local approximation also requires that $|m|/\varpi \ll k$; we set $|m|\leq 15$ to ensure that the local approximation is satisfied even for the largest-scale modes with $k = 2\pi/H_{\rm P}$; the maximum value of $H_{\rm P}$ in the tachocline and convective region is always smaller than $0.1r_\odot$. We investigate the growth rates of small-scale modes with larger toroidal wavenumbers in Section 4.3.3. Because the final wavenumber parameter k_{ϖ}/k_Z is not constrained by our approximations, we vary it freely, including both positive and negative values with magnitudes greater than or less than 1.

We next choose an initial magnetic field strength and geometry. We assume a small poloidal magnetic field of magnitude 0.2G and explore the effects of varying this magnitude in Section 4.3.2.5. To explore the effects of the poloidal magnetic field geometry, we set B_{ϖ}/B_Z corresponding to a field oriented in the r, θ, ϖ, and Z directions; reversing the direction of this field is equivalent to making the substitution $m \to -m$, so it is unnecessary to consider the opposite orientations. We determine the strength of the toroidal magnetic field at a given point by setting $R_{\mathrm{TP}} = 5$, consistent with the expected dominance of the toroidal magnetic field in stellar MHD equilibria (Braithwaite, 2009); we investigate the effects of varying R_{TP} in Section 4.3.3.2. Setting R_{TP} fixes the values of $\omega_{\mathrm{A}\phi}$ and $(\mathbf{k} \cdot \mathbf{v}_{\mathrm{A}})_{\phi}$. Having set the ranges of the parameters, we then calculate the nonaxisymmetric growth rate $\Gamma = -i\omega$ for $\sim 10^6$ wavevectors in the phase space, and compare the results for all of the indicated field geometries at each location in the Sun.

4.3 Results

We now discuss our calculations of the growth rates of modes throughout the Sun. In Section 4.3.1, we present the growth rates of the fastest-growing modes of the overall instability in the Sun, and discuss whether shear or convection is responsible for driving the instability at each location. In Section 4.3.2 we discuss the nature of the axisymmetric instability throughout the Sun, and identify those modes that are sensitive to the initial magnetic field geometry. Finally, in Section 4.3.3, we discuss nonaxisymmetric effects and their variation with the toroidal field and the initial poloidal magnetic field strength.

4.3.1 Growth Rates of Instability in the Sun

Figure 4.1 shows the maximum growth rate, Γ, of the unstable modes of the dispersion relation Equation (4.16) at each location in units of the local angular rotation velocity Ω at each location in the Sun; for this calculation, the maximum growth rate is calculated for any poloidal field geometry. It is clear that the instability grows quickly throughout the tachocline and the solar convective region. The Sun may be divided into four regions in which the characteristics of the fastest-growing modes have significant differences. Region TS is located in the stably stratified part of the tachocline at $r < 0.713r_\odot$. The fastest-growing modes in this region are nonaxisymmetric, with the maximum growth rate corresponding to the largest magnitude for the toroidal wavenumber $|m|$; we discuss nonaxisymmetric effects in Section 4.3.3. Region TU is located in the convectively unstable region close to the tachocline at $r > 0.713r_\odot$, and has a colatitude range of $0^o < \theta < 60^o$ that is similar to but slightly larger than that for Region TS, and corresponds closely to the region for which the radial shear in the tachocline is negative. In both Region TS and Region TU, the fastest growing modes have growth rate Γ on the order of Ω, although the growth rates tend to be significantly lower very close to the poles and for $\theta > 45^o$. Note that the somewhat smaller growth rates in Region TS result from the fact that the GONG measurement of Ω in this region is taken at $r = 0.692r_\odot$, which is at the bottom of the tachocline; a calculation using a slightly smaller value for N^2 corresponding to $r \sim 0.70r_\odot$ gives similar growth rates to those found in Region TU. The growth rate in both Regions TS and TU is maximized at $\theta \sim 20^{\rm o}$–$30^{\rm o}$, which is the approximate location in the tachocline where the shear in the ϖ direction is largest. This indicates that the strong shear in the tachocline is probably driving the growth of instability. The lack of growth of modes in Region TS for colatitudes $53^{\rm o} < \theta < 60^{\rm o}$ is probably a result of the strong stable stratification in this region, where the radial shear is relatively weak.

Region TL is located at lower latitudes corresponding to $\theta > 60^{\rm o}$ in the convectively unstable region $r > 0.713r_\odot$ of the tachocline and lower convective zone. The typical growth rate of instabilities in Region TL is typically much smaller than both the rotation rate Ω and the magnitude of the buoyancy frequency $|N|$; this is likely because the positive radial shear retards the growth of instability. Finally, in Region C, located at $r > 0.8r_\odot$ at all latitudes, the typical growth rate of the fastest growing modes is similar to the local magnitude of the buoyancy frequency $|N|$. This indicates that the modes in this region are driven by convection.

In order to more precisely determine whether convection or shear is responsible for the growth of instability in each location, we repeat the calculation of the growth rate with only convection present (by setting all derivatives of the rotation rate Ω to 0) and with only shear present (by setting the buoyancy frequency $N = 0$) and compare the resulting growth rates. Figure 4.2 shows those locations where the growth of instability is driven by convection and by shear. The figure indicates that shear is the only mechanism that can drive instability in the stably stratified Region

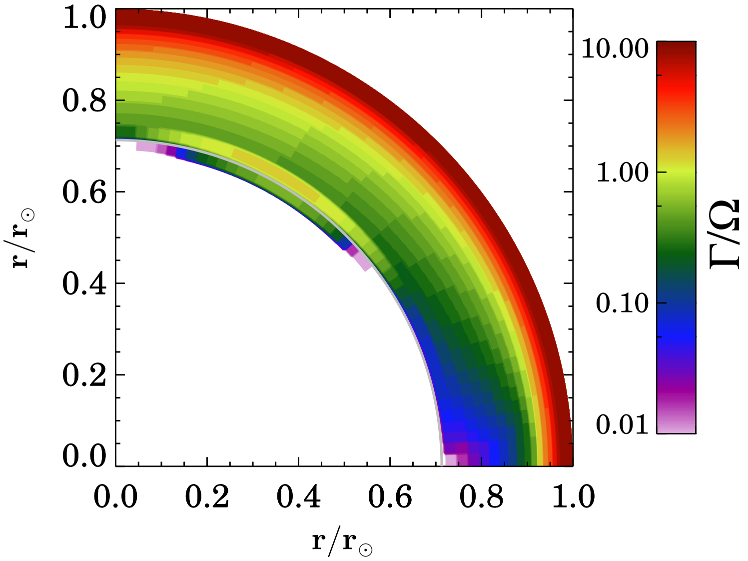

Figure 4.1 The growth rate Γ of the instability throughout the Sun in terms of the local rotation rate Ω. Black lines represent contours of Ω. The gray line indicates the bottom of the convection zone in the tachocline. For definiteness, we do not indicate that growing modes are present unless their growth rate is larger than 0.01Ω. The rapid growth of this instability throughout the convective zone indicates that it may play a role in the origin of the solar magnetic field.

TS, as expected from the analysis in Section 4.1.2.1. At the bottom of Region TU, shear is the dominant driver of instability, in agreement with our previous conclusions. The driver of modes at the bottom of Region TL at lower latitudes varies; modes are driven by convection at $60^o < \theta < 73^o$, and by shear for $73^o < \theta < 90^o$. We will discuss the cause of this variation in Section 4.3.2.3. In the rest of the convectively unstable region of the Sun, including the upper parts of Region TL and TU and the entirety of Region C, convection is responsible for driving the growth of instability.

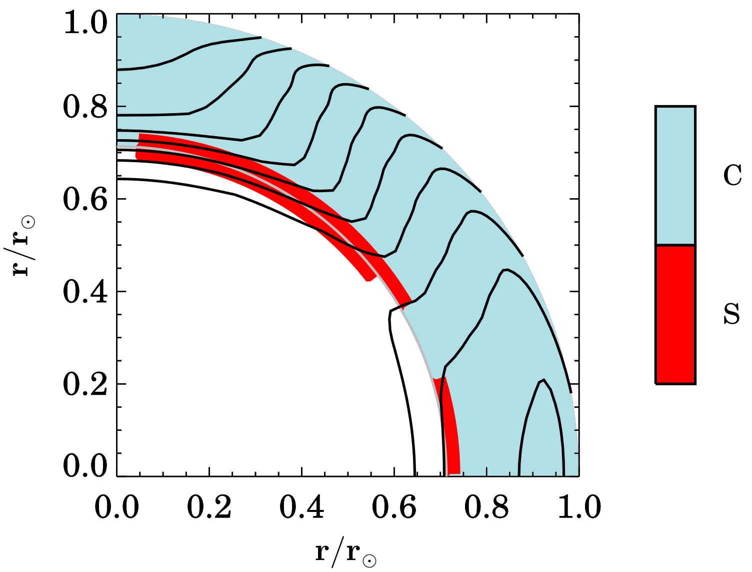

Figure 4.2 The locations in the Sun where the growth of instability is dominated by convection (shown in blue) and shear (shown in red). The black lines are contours of Ω found by GONG; they show the layers of strong shear in the tachocline and near the solar surface. Shear is dominant at most latitudes close to the tachocline, while convection dominates near the tachocline at $\theta \sim 60^{\rm o}$ and throughout the rest of the convective zone.

4.3.2 Axisymmetric Mode Analysis

In this section, we discuss the variation of the growth rate Γ with initial parameters for axisymmetric modes with $m = 0$. We find that for axisymmetric modes, there is no significant dependence of growth rate on the sign of $\mathbf{k}\cdot\mathbf{v}_{\rm A}$; therefore, in our analysis of variation with wavenumber we present the variation of growth rate with k_{ϖ}/k_Z and $|\mathbf{k}\cdot\mathbf{v}_{\rm A}|$.

Our analysis of the variation of growth rate with the wavenumber is divided into four parts, corresponding to the four regions of the Sun in which growth of instability can occur discussed in the previous section: Regions TS, C, TL, and TU. We choose an initial magnetic field geometry with $|B_{\text{pol}}| = 0.2G$, $B_\varpi/B_Z = \tan\theta$, which corresponds to a magnetic field oriented in the r direction, and a toroidal to poloidal field ratio of $R_{\text{TP}} = 5$.

4.3.2.1 Stably Stratified Region TS

In the stably stratified tachocline corresponding to Region TS, only shear can drive unstable modes. While the radial shear $d\Omega^2/d\ln\varpi$ is negative for $\theta < 62^{\text{o}}$, we observe modes with significant growth only for $\theta < 53^{\text{o}}$.

Figure 4.3 shows the variation with k_ϖ/k_Z and $|\mathbf{k}\cdot\mathbf{v}_{\text{A}}|$ of the growth rate Γ at $r = 0.692r_\odot$, $\theta = 30^{\text{o}}$ in Region TS. It clearly shows two major regions of instability: first, a region at low $|k_\varpi/k_Z| \ll 1$ and $|\mathbf{k}\cdot\mathbf{v}_{\text{A}}| \approx 0.5\Omega$, and second, a region at $k_\varpi/k_Z \approx -1.57$ with a wider variation of $|\mathbf{k}\cdot\mathbf{v}_{\text{A}}|$. The first region corresponds to nearly vertical modes, which do not couple to shear in the Z direction or to the magnetic field in the ϖ direction. These modes therefore couple to the star as though it were cylindrically symmetric, and are thus analogous to the classical MRI presented by Balbus & Hawley (1991) in accretion disks; the only difference is that the shear profile is non-Keplerian and buoyant effects in the ϖ direction are present. We therefore call this the classical MRI (CMRI) mode. The fastest growth rate in this region is $\Gamma = 0.073\Omega$, and it is located at $k_\varpi/k_Z = 0.16$, $|\mathbf{k}\cdot\mathbf{v}_{\text{A}}| = 0.83\Omega$.

The second region of phase space where fast growth occurs corresponds to modes with wavenumber nearly perpendicular to the magnetic field, so that $|\mathbf{k}\cdot\mathbf{v}_{\text{A}}| \ll kv_{\text{A}}$; we call these modes perpendicular small scale (PSS) modes. PSS

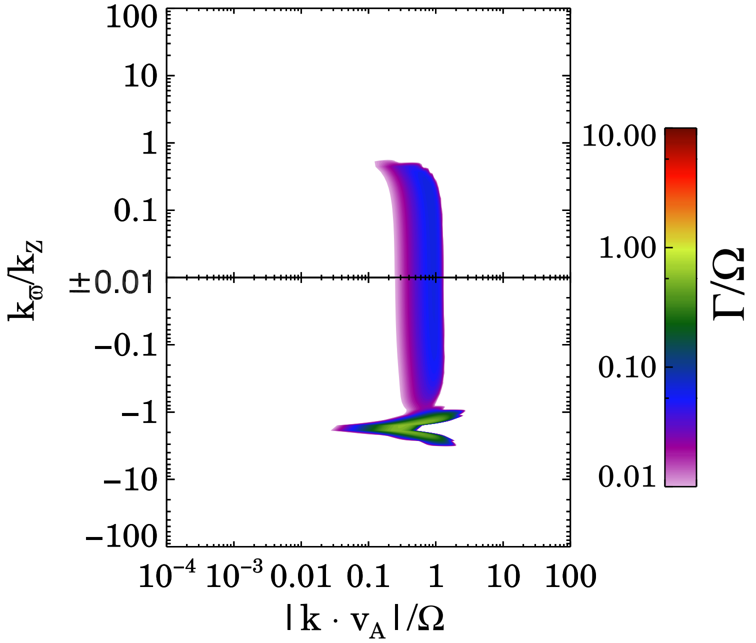

Figure 4.3 The dependence of the growth rate Γ on the phase space parameters k_ϖ/k_Z and $|\mathbf{k}\cdot\mathbf{v}_\mathrm{A}|$ at the location $r = 0.692r_\odot$, $\theta = 30^\mathrm{o}$ in Region TS; the growth rate and $|\mathbf{k}\cdot\mathbf{v}_\mathrm{A}|$ are given in units of the local rotation rate Ω. The figure shows two regions of fast growth; the first has $k_\varpi/k_Z \ll 1$ and $|\mathbf{k}\cdot\mathbf{v}_\mathrm{A}| \approx 0.5-1.0\Omega$, while the second has $k_\varpi/k_Z \approx -1.5$ and a wide range of $|\mathbf{k}\cdot\mathbf{v}_\mathrm{A}|$ $\approx$0.03Ω–2.0Ω. Note that the two "‘wings"’ of the PSS mode correspond to opposite signs of $(\mathbf{k}\cdot\mathbf{v}_\mathrm{A})_\mathrm{pol}$. The fastest-growing mode has a growth rate $\Gamma = 0.63\Omega$, and is located in the second region at the coordinates $k_\varpi/k_Z = -1.57$, $|\mathbf{k}\cdot\mathbf{v}_\mathrm{A}| = 0.32\Omega$.

modes have large k which greatly weakens the stable stratification, but the moderate value for $\mathbf{k}\cdot\mathbf{v}_\mathrm{A}$ means that the magnetic tension does not greatly reduce the growth rate. The growth of these short-wavelength modes is primarily inhibited by resistive dissipation, because of the large k required to reduce the stable stratification by such a large factor; typically, $k^2\eta \sim \Omega$. Therefore, strongly negative radial shear is required to drive these modes. Note that for very precise orientations of k such that $|\mathbf{k}\cdot\mathbf{v}_\mathrm{A}| \ll 0.001\Omega$ essentially hydrodynamic modes with similar k may exist;

however, these modes do not grow at a significantly faster rate than PSS modes and they represent a very small portion of the phase space. As we show in Section 4.3.2.5, however, these hydrodynamic modes become more important for smaller initial field magnitudes. Both the PSS mode and the CMRI mode are small-scale shear modes which may be properly referred to as MRI modes.

For the radial magnetic field geometry, this region of phase space only has growing modes at colatitudes $9^{\rm o} < \theta < 45^{\rm o}$, which correspond to shears of $d\ln\Omega/d\ln\varpi < -0.25$. We find that for $r = 0.692r_\odot$, $\theta = 30^{\rm o}$, the fastest-growing mode in this region of phase space has a growth rate $\Gamma = 0.63\Omega$, and is located at the coordinates $k_\varpi/k_Z \approx -1.57$, $|\mathbf{k}\cdot\mathbf{v}_{\rm A}| = 0.32\Omega$. At latitudes where field-parallel modes in the second region are present, their growth rate is larger than that of nearly vertical modes in the first region by a factor of ~ 10; at $r = 0.692r_\odot$, $\theta = 30^{\rm o}$, the ratio of the growth rates is approximately 8.6.

4.3.2.2 Convectively Dominated Region C

In region C, growing modes are dominated by convection, while shear plays little role in the growth of instability, especially near the top of the convection zone. In Figure 4.4, we show the dependence of the growth rate Γ of the dispersion relation (4.16) on the parameters k_ϖ/k_Z and $|\mathbf{k}\cdot\mathbf{v}_{\rm A}|$ at the location $r = .965r_\odot$, $\theta = 37.5^{\rm o}$ in the upper part of Region C. We find that the fastest-growing modes at this location are those that have $k_\varpi/k_Z = -1.31$ and $|\mathbf{k}\cdot\mathbf{v}_{\rm A}| = 7.8\times10^{-6}\ \Omega$, with maximum growth rate $\Gamma = 10.4\Omega$, similar to the local magnitude of the buoyancy frequency $|N| = 10.5\Omega$. The dependence of the growth rate on $|\mathbf{k}\cdot\mathbf{v}_{\rm A}|$ is very weak over most of the phase space, with a slight decrease in growth rate at smaller scales; however, at values of $\mathbf{k}\cdot\mathbf{v}_\mathrm{A}$ corresponding to $k^2\xi \approx N$, the growth rate abruptly drops from $\Gamma \approx N$ to $\Gamma = 0$ as one moves in the direction of higher wavenumber. Note that the drop in the value of $\mathbf{k}\cdot\mathbf{v}_\mathrm{A}$ corresponding to this cutoff at $k_\varpi/k_Z \sim -1.3$ occurs because this orientation of the wavenumber corresponds to $|\mathbf{k}\cdot\mathbf{v}_\mathrm{A}| \ll kv_\mathrm{A}$; the drop still occurs at $k^2\xi \sim N$. The variation of growth rate with k_ϖ/k_Z is also relatively simple; there is a region of very low growth at $k_\varpi/k_Z \sim 0.75$, while a region of fast growth is observable for $k_\varpi/k_Z \approx -1.3$, which is the approximate location of the fastest-growing mode

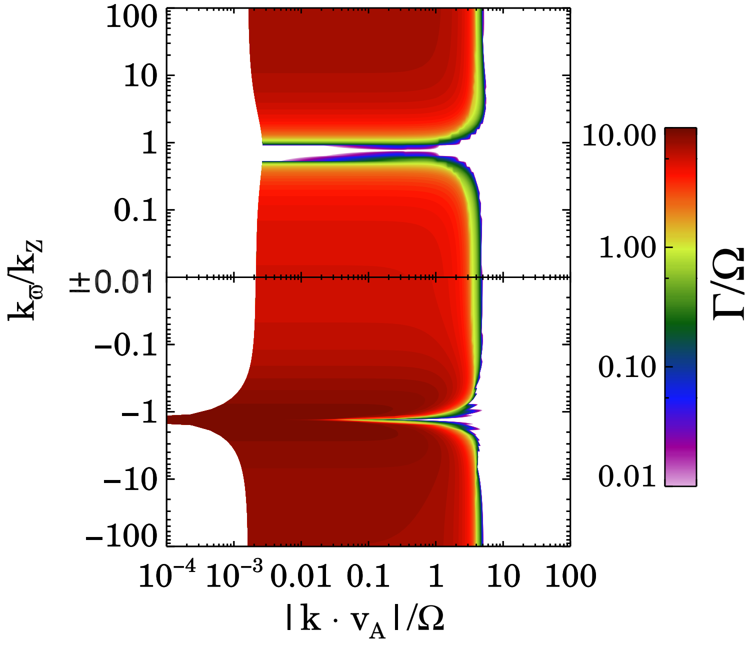

Figure 4.4 The dependence of the growth rate Γ on the phase space parameters k_ϖ/k_Z and $|\mathbf{k}\cdot\mathbf{v}_\mathrm{A}|$ at the location $r = .965r_\odot$, $\theta = 37.5^\circ$ in Region C; the growth rate and $|\mathbf{k}\cdot\mathbf{v}_\mathrm{A}|$ are given in units of the local rotation rate Ω. The fastest growth is found at $k_\varpi/k_Z = -1.31$ and $|\mathbf{k}\cdot\mathbf{v}_\mathrm{A}| = 7.8\times10^{-6}\Omega$, which corresponds to a hydrodynamic convective mode.

mentioned above. Because this variation is most evident for modes that are very large in scale, we may identify the causes of this variation by using the dispersion relation Equation (4.25) for such modes. For $r = .965r_{\odot}$, $\theta = 37.5^{\mathrm{o}}$, which is near the outermost layers of Region C, the appropriate regime is $|N| \gg \kappa, \Omega$. In this case, the growth rate $\Gamma_{\mathrm{conv}} = -i\omega$ of large scale convective modes is given by

$$\Gamma_{\mathrm{conv}} \approx \frac{1}{\sqrt{1+(k_{\varpi}/k_Z)^2}} \left| (\frac{k_{\varpi}}{k_Z} \cos\theta - \sin\theta) N \right|, \tag{4.49}$$

The form of Equation (4.49) indicates that in general, we expect modes to grow faster for $k_{\varpi}/k_Z < 0$, which is indeed what we observe.

To understand the regions of fastest and slowest growth, we calculate the extrema of the growth rate given by Equation (4.49). One extremum of the growth rate Γ_{conv} occurs at $k_{\varpi}/k_Z = \tan\theta$, resulting in $\Gamma_{\mathrm{conv}} = 0$. For $r = .965r_{\odot}$, $\theta = 37.5^{\mathrm{o}}$, this extremum corresponds to $k_{\varpi}/k_Z = 0.77$, so it explains the region of low growth that we observe at $k_{\varpi}/k_Z \approx 0.75$. The second extremum of the growth rate occurs at $k_{\varpi}/k_Z = -\cot\ \theta$, which results in $\Gamma_{\mathrm{conv}} = |N|$. For the location discussed in this section, $\cot\theta = -1.31$; therefore, this analysis successfully predicts the orientation of the poloidal wavevector for the fastest-growing mode. Note that the growth rate Γ found in our full analysis is typically smaller than N because for $k_{\varpi}/k_Z = -\cot\theta$, $\widetilde{\kappa} > 0$.

To understand the physical reasons for the observed dependence of growth rate on k_{ϖ}/k_Z, we use Equation (4.7) to calculate the orientation of the perturbed velocity flows for the two extrema. At $k_{\varpi}/k_Z = \tan\theta$, the perturbed flows have the ratio $\delta v_{\varpi}/\delta v_Z = -\cot\theta$, which means that the perturbed velocity field is oriented in the $\pm\theta$ direction. For $k_{\varpi}/k_Z = -\cot\theta$, we find that $\delta v_{\varpi}/\delta v_z = \tan\theta$, which means that the perturbed flows are oriented in the $\pm r$ direction. Because convective instability drives flows in the r direction, we expect that the growth rate of convective modes will be $\sim N$ for $k_{\varpi}/k_Z = -\cot\theta$, and ~ 0 for $k_{\varpi}/k_Z = \tan\theta$, and this is exactly what we observe in our analysis of the dispersion relation.

In this analysis, we have focused on the outermost part of Region C, where convection is completely dominant; however, for $r < 0.9r_{\odot}$ there are parts of phase space for which $k^2\xi > N$ and $|\mathbf{k}\cdot\mathbf{v}_{\rm A}| < \Omega$, in which convective effects are negligible and only shear can drive modes. These small-scale modes may have important effects on the nonlinear evolution of instability, even though they grow more slowly than the convective large-scale modes discussed here.

4.3.2.3 Region TL

In Region TL, the radial shear is positive, which typically inhibits the growth of modes. Nevertheless, modes can driven by either shear or convection in this region, as shown earlier in Section 4.3.1. Figure 4.5 shows the dependence of the growth rate on the phase space parameters k_{ϖ}/k_Z and $|\mathbf{k}\cdot\mathbf{v}_{\rm A}|$ at the location $r = .721r_{\odot}$, $\theta = 75^{\rm o}$ in Region TL, which is a location where the growth rates of shear modes and convective modes are approximately equal. We observe unstable modes only for $|k_{\varpi}/k_Z| > 10$, which means that the unstable modes are oriented nearly in the $\pm\varpi$ direction, which corresponds to unstable displacements oriented nearly in the $\pm Z$ direction. Because $|\mathbf{k}\cdot\mathbf{v}_{\rm A}| < 0.01\Omega$ for all growing modes, the magnetic field has no significant effect in this region, and all modes are large-scale.

The maximum growth rate we observe at this location is $\Gamma = 0.096\Omega$, which corresponds to the parameters $k_\varpi/k_Z = 33.8$ and $|\mathbf{k}\cdot\mathbf{v}_\mathrm{A}| = 5.27\times10^{-5}\ \Omega$; this is significantly smaller than the local values of $|N|$ and Ω. Comparison of the left and middle panels of Figure 4.5 reveals that the presence of positive radial shear significantly reduces the growth rate of convective modes with negative k_ϖ/k_Z, while producing growing modes with positive k_ϖ/k_Z that, for this location, grow slightly faster than convective modes. It is important to note that because the radius of the strongly sheared region known as the tachocline decreases one moves to lower latitudes, it is possible that no part of the tachocline is in the convectively unstable region with $r > 0.713r_\odot$ at latitudes corresponding to Region TL (Basu & Antia, 2001). If so, the effects of convection will be dominant throughout Region TL, shear will be significantly smaller at the bottom of Region TL, but the typical fastest-growing modes will have similar growth rates to those calculated in this section. Because all growing modes are large-scale, we can again make use of Equation (4.25) to detail their properties. The growth rate $\Gamma_\mathrm{shear} \equiv -i\omega$ of modes driven by shear in the tachocline can then be found by setting $N = 0$ in Equation (4.25). In the tachocline, the shear is oriented approximately in the spherical r direction, so

$$\widetilde{\kappa}^2 \approx 2\Omega^2\left(2+q(\sin^2\theta-\frac{k_\varpi}{k_Z}\cos\theta\sin\theta)\right), \tag{4.50}$$

where $q \equiv d\ln\Omega/d\ln r$. Therefore, the growth rate of shear modes is given by

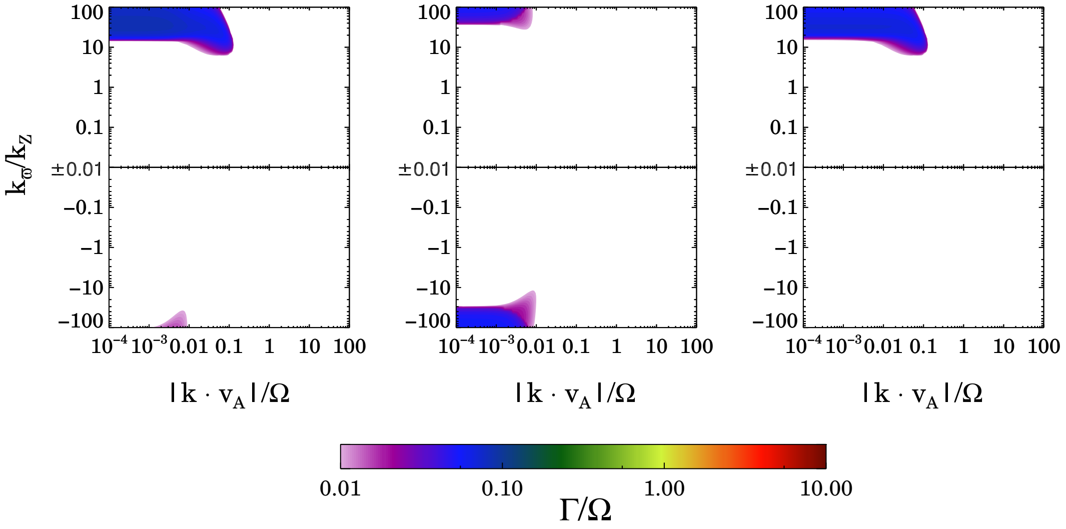

Figure 4.5 The dependence of the growth rate Γ on the phase space parameters k_ϖ/k_Z and $|\mathbf{k}\cdot\mathbf{v}_A|$ at the location $r = .721 r_\odot$, $\theta = 75^\circ$ in Region TL including (left panel) the effects of both convection and shear, (middle panel) the effects of convection only, setting gradients of Ω to 0, (right panel) the effects of shear only, setting $N = 0$. In all cases, the growth rate is given in units of the local rotation rate Ω. The fastest growth rate of unstable modes, which is significantly smaller than that in the other regions of the Sun, is found at $k_\varpi/k_Z = 33.8$ and $|\mathbf{k}\cdot\mathbf{v}_A| = 5.27\times10^{-5}\ \Omega$.

$$\Gamma_{\text{shear}} \approx \Omega\sqrt{-\frac{4+2q(\sin^2\theta - k_\varpi/k_Z\cos\theta\sin\theta)}{1+(k_\varpi/k_Z)^2}}. \tag{4.51}$$

In Region TL, $q > 0$, so growing modes exist only if

$$\frac{k_\varpi}{k_Z} > \tan\theta + \frac{2}{q\sin\theta\cos\theta}. \tag{4.52}$$

The growth rate of large-scale convective modes in Region TL may be calculated by setting gradients of Ω to 0 in Equation (4.25). The resulting growth rate $\Gamma_{\text{conv}} = -i\omega$ is then given by

$$\Gamma_{\rm conv} \approx \sqrt{-\frac{4\Omega^2 + ([k_\varpi/k_Z]\cos\theta - \sin\theta)^2 N^2}{1+(k_\varpi/k_Z)^2}}. \tag{4.53}$$

Growing convective modes exist only for

$$\left|\frac{k_\varpi}{k_Z} - \tan\theta\right| > \frac{2\Omega}{|N|\cos\theta}. \tag{4.54}$$

Neither type of mode can grow for $|k_\varpi/k_Z| < \tan\theta$; this explains why growing modes do not exist except at large magnitudes of $|k_\varpi/k_Z|$ in Region TL, where $\tan\theta > 1$. At the location $r = .721 r_\odot$, $\theta = 75^{\rm o}$, the dimensionless shear is given by $q = 1.025$ and the local buoyancy frequency has magnitude $|N| = 0.244\Omega$. Equation (4.52) therefore predicts that growing shear modes must have $k_\varpi/k_Z > 11.5$, while Equation (4.54) predicts that $k_\varpi/k_Z > 35.4$ or $k_\varpi/k_Z < -27.9$. The right and middle panels of Figure 4.5 reveal that these conditions are indeed obeyed in the full analysis of shear and convective modes.

In order to understand why the driver of modes varies with latitude in Region TL, we now consider the variation of the growth rates of these modes with latitude. In the limit $|k_\varpi/k_Z| \gg 1$ applicable to modes in Region TL, the convective growth rate given by Equation (4.53) becomes $\Gamma_{\rm conv} \approx |N|\cos\theta$; therefore, the growth rate of convective modes decreases sharply as θ increases. In contrast, the growth rate of shear modes in Region TL does not vary sharply with θ, although it does increase slightly with θ over most of region TL because both the radial shear q and the rotational frequency Ω increase with θ. Near the equator, however, the growth rate of shear modes decreases again, because less and less of the radial shear is oriented in the Z direction, reducing the destabilization resulting from the condition given in

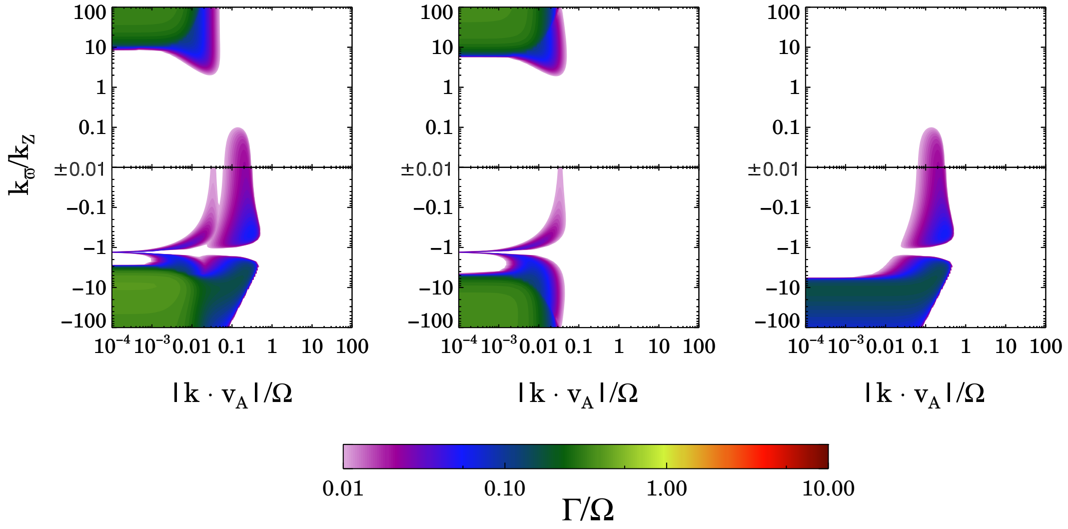

Figure 4.6 The dependence of the growth rate Γ on the phase space parameters k_ϖ/k_Z and $|\mathbf{k}\cdot\mathbf{v}_\mathrm{A}|$ at the location $r = .749r_\odot$, $\theta = 37.5^\mathrm{o}$ in Region TU including (left panel) the effects of both convection and shear, (middle panel) the effects of convection only, with gradients of Ω set to 0, (right panel) the effects of shear only, setting $N = 0$. In all cases, the growth rate is given in units of the local rotation rate Ω. The figure's middle and right panels reveal the presence of two shear modes and two convective modes, which are discussed in the text. The fastest growth rate is $\Gamma = 0.43\Omega$ at the coordinates $k_\varpi/k_Z = -9.56$ and $|\mathbf{k}\cdot\mathbf{v}_\mathrm{A}| = 4.0\times10^{-5}\ \Omega$; the growth is driven by both convection and shear.

Equation (4.43). As a result, the overall growth rate falls below 0.05Ω for $\theta > 83^\mathrm{o}$. Because of the observed suppression of all growing modes near the equator of the Sun, we expect that the production of large-scale features will be inhibited near the equator, especially at $\theta > 83^\mathrm{o}$; we discuss the implications of this result for solar activity in Section 4.4.2.

4.3.2.4 Region TU

In Region TU the variety of growing modes is greater than in other regions. There are two important cases where the growing modes are significantly different. In the upper tachocline at $r = 0.721 r_\odot$, the shear is very strong, $|q| \gg 1$, while the buoyancy frequency is small compared to the rotation frequency $N \ll \Omega$. Thus, we expect that modes will be driven by shear. At slightly larger radii, shear and convection both contribute significantly to the growth of modes, and both are of the same order as Ω. We will first discuss the case where both convection and shear are important, and then discuss how the results change in the shear-dominated case.

Figure 4.6 shows the dependence of the growth rate Γ of the dispersion relation (4.16) on the parameters k_ϖ/k_Z and $|\mathbf{k}\cdot\mathbf{v}_\mathrm{A}|$ at the location $r = 0.749 r_\odot$, $\theta = 37.5^\mathrm{o}$. The left panel shows the growth rate including the effects of both shear and convection, while the middle and right panels show the growth rate including only convection (setting gradients of Ω to 0), and including only shear (setting $N = 0$). The figure reveals that there are two types of modes driven by shear and two types of modes driven by convection in Region TU. The two types of shear mode shown in the right panel are a large-scale hydrodynamic mode with a wide range of values of $|\mathbf{k}\cdot\mathbf{v}_\mathrm{A}|$ and $k_\varpi/k_Z < -10$, and a small-scale MHD mode with $0.025\Omega < |\mathbf{k}\cdot\mathbf{v}_\mathrm{A}| < 0.5\Omega$ and $-1.0 < k_\varpi/k_Z < 0.1$. The two types of convective mode shown in the middle panel are a mode with large $|k_\varpi/k_Z|$ and $k^2\xi < N$ and a highly overstable mode that corresponds to $k^2\xi \approx N$ and exists at all values of $k_\varpi/k_Z < 0$. The two large-scale modes are very similar to the convective and shear modes discussed in Section 4.3.2.3; the sole difference is that in Region TU, $q < 0$, so the large-scale shear mode grows only for $k_\varpi/k_Z < 0$ and inhibits growth for $k_\varpi/k_Z > 0$. The fastest growth rate found in the full analysis is $\Gamma = 0.432\Omega$, which occurs at the coordinates $k_\varpi/k_Z = -9.58$ and $|\mathbf{k}\cdot\mathbf{v}_\mathrm{A}| = 4.0 \times 10^{-5}\ \Omega$. The instability at this location is driven by both convection and shear; the growth rate including only convection is 0.339Ω, while the growth rate including only shear is 0.168Ω.

The small-scale shear mode in Region TU is similar to the CMRI mode discussed in Section 4.3.2.1, and it can be referred to as an MRI mode. In Region TU it exists for all $\theta < 60^\mathrm{o}$, which corresponds to the region where $d\ln\Omega/d\ln\varpi < 0$. The peak growth rate for this mode at the location $r = .749r_\odot$, $\theta = 37.5^\mathrm{o}$ is $\Gamma = 0.120\Omega$, and occurs at the coordinates $k_\varpi/k_Z = -9.56$ and $|\mathbf{k}\cdot\mathbf{v}_\mathrm{A}| = 4.0 \times 10^{-5}\ \Omega$. The overstable convective mode at $k^2\xi = N$ has maximum growth rate $\Gamma = 0.048\Omega$, while its oscillatory frequency is $\approx 1.05\Omega$.

We now discuss the growth rate of modes in the upper tachocline, where shear is dominant. Figure 4.7 shows the dependence of the growth rate Γ of the dispersion relation 4.16 on the parameters k_ϖ/k_Z and $|\mathbf{k}\cdot\mathbf{v}_\mathrm{A}|$ at the location $r = 0.721r_\odot$, $\theta = 37.5^\mathrm{o}$. In this figure, the two shear modes have merged, and their growth rate is significantly larger due to the strong shear in the tachocline. The convective modes have been swamped by the shear modes, except at $k_\varpi/k_Z > 10$, where a convective mode is present but strongly suppressed compared to the growth rate that it would have in the absence of shear.

The figure shows that the fastest growing mode occurs for parameters $k_\varpi/k_Z = -1.202$, $|\mathbf{k}\cdot\mathbf{v}_\mathrm{A}| = 2.26 \times 10^{-6}\Omega$. These parameters correspond to a large-scale hydrodynamic shear mode; because the effect of convection is negligible compared

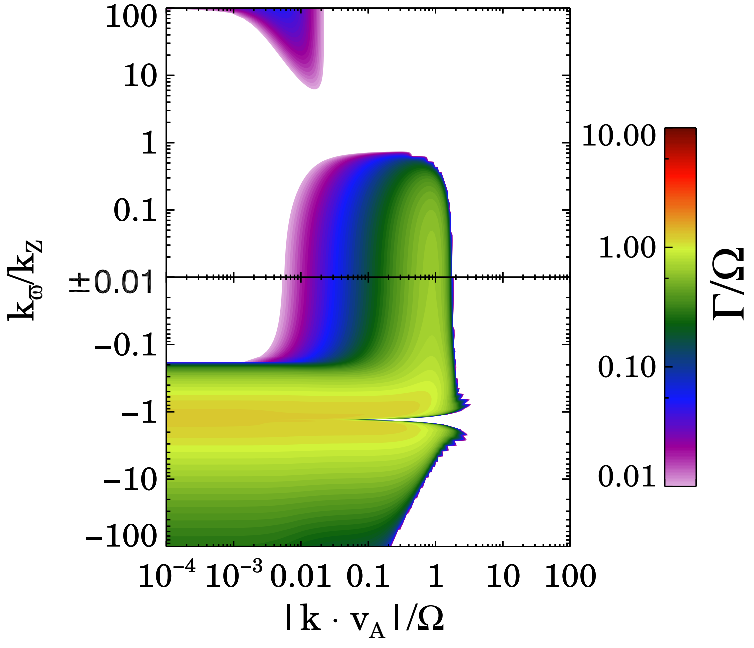

Figure 4.7 The dependence of the growth rate Γ on the phase space parameters k_ϖ/k_Z and $|\mathbf{k}\cdot\mathbf{v}_\mathrm{A}|$ at the location $r = 0.721r_\odot$, $\theta = 37.5^\circ$ in Region TU; the growth rate is given in units of the local rotation rate Ω. The fastest growth is found at $k_\varpi/k_Z = -1.202$ and $|\mathbf{k}\cdot\mathbf{v}_\mathrm{A}| = 2.26\times10^{-6}\Omega$, which corresponds to a hydrodynamic shear mode.

to that of shear, its growth rate is

$$\Gamma_\mathrm{shear} \approx \Omega\sqrt{-\frac{4+2q(\sin^2\theta-[k_\varpi/k_Z]\cos\theta\sin\theta)}{1+(k_\varpi/k_Z)^2}}, \tag{4.55}$$

For our location at $r = 0.721r_\odot$, $\theta = 37.5^\circ$, $q = -4.446$. The only extremum that corresponds to a growing mode occurs at $k_\varpi/k_Z = -1.18$, which gives $\Gamma_\mathrm{shear} = 1.35\Omega$, which are close to the fastest-growing mode location $k_\varpi/k_Z = -1.20$ and growth rate $\Gamma = 1.38$ calculated in the full analysis. It is important to note that the small-scale shear mode at low $|k_\varpi/k_Z|$ grows nearly as fast as does the large-scale shear mode in the upper tachocline. Thus, the small-scale MHD shear modes are likely to be important in both Region TU and Region TS. These modes, which have $\mathbf{k} \cdot \mathbf{v}_{\mathrm{A}} \sim \Omega$ and are driven by shear, can properly be called MRI modes due to their similarity to the simpler modes found in accretion disks.

4.3.2.5 Effect of Field Geometry on Axisymmetric Modes

We find that neither hydrodynamic shear modes nor convectively driven modes are strongly affected by the field geometry, except in the case, unrealistic for the Sun, where the magnitude of the magnetic field is large enough that it becomes dynamically important. The apparent dependence on $\mathbf{k} \cdot \mathbf{v}_{\mathrm{A}}$ for such modes is merely a dependence on k. In contrast, the small-scale shear modes found in Regions TS and TU, which are true MRI modes, have growth rates that depend on the field geometry. For small-scale modes, $\omega_{\mathrm{A}\phi} \ll \mathbf{k} \cdot \mathbf{v}_{\mathrm{A}}$, so the growth rates depend only on the poloidal magnetic field ratio B_{ϖ}/B_Z and the magnitude of the poloidal field B_{pol}. We now calculate maximum growth rates for each type of mode in Regions TS and TU. We choose values of B_{ϖ}/B_Z corresponding to a field oriented in the r, θ, and Z directions[2] and field magnitudes ranging from 10^{-4} G to 10^{4} G. In order to isolate the CMRI mode and the small-scale shear mode in Region TU, which can overlap with other modes, we calculate their maximum growth rates only for the region of phase space with $|k_{\varpi}/k_Z| < 0.1$.

The first type of small-scale shear mode is the CMRI mode in Region TS, which typically corresponds to small $|k_{\varpi}/k_Z|$. At the location $r = 0.692 r_{\odot}$, $\theta = 30^{\mathrm{o}}$,

---

[2] We do not carry out an detailed analysis using fields oriented in the ϖ direction because the CMRI and PSS modes overlap with eachother for this field orientation; this significantly complicates the analysis.

we find that for a poloidal field magnitude of 0.2 G the maximum growth rates for this type of mode are 0.073Ω, 0.142Ω, and 0.063Ω for B oriented in the r, θ, and Z directions, respectively. CMRI modes with growth rate larger than 0.01Ω exist for the CMRI mode only in the narrow range of poloidal field magnitudes 0.02 G$< B_{\rm pol} < 0.6$ G. For larger magnetic fields, all modes with k large enough to reduce the effects of stratification have $|\mathbf{k}\cdot\mathbf{v}_{\rm A}| \gg \Omega$; as a result, the magnetic tension prevents the growth of instability. In contrast, for smaller fields the condition $\mathbf{k}\cdot\mathbf{v}_{\rm A} \sim \Omega$ for small-scale shear modes implies that k is very large; the growth of these modes is then strongly inhibited by resistive and viscous dissipation, because the characteristic dissipative frequencies are proportional to k^2.

The second type of mode found in Region TS is the PSS mode with $|\mathbf{k}\cdot\mathbf{v}_{\rm A}| \ll kv_A$; for PSS modes the stabilizing effects of magnetic tension are substantially reduced. At the location $r = 0.692r_\odot$, $\theta = 30^{\rm o}$, this mode grows quickly only for relatively large poloidal field strengths 0.08 G$< B_{\rm pol}$, and the growth rate depends strongly on the orientation of the field. For a poloidal field of 0.2 G the growth rate of this mode is 0.63Ω for a magnetic field oriented in the r direction, while the growth rate for modes oriented in the θ and Z directions is smaller than 0.1Ω and represents an extension of the CMRI mode. Figure 4.8 shows the dependence of the growth rate on the phase space parameters k_ϖ/k_Z and $|\mathbf{k}\cdot\mathbf{v}_{\rm A}|$ at the location $r = 0.692r_\odot$, $\theta = 30^{\rm o}$ for a magnetic field oriented in the θ direction, revealing the presence of the CMRI mode and the absence of the PSS mode. Comparing the growth rates for various field orientations B_ϖ/B_Z indicates that the fastest growth occurs for field orientations that are close to radial, but shifted towards $B_\varpi/B_Z = 1$. At $r = 0.692r_\odot$,

$\theta = 30^{\text{o}}$, the PSS mode grows most quickly for $B_{\varpi}/B_Z = 0.84$; growth rates are significantly larger than those for the CMRI mode for field orientations within $\pm 45^{\text{o}}$ of this fastest-growing orientation. The range of orientations is typically similar at other latitudes. For larger $B_{\text{pol}} \gg 1$G, the PSS mode grows at similar rates to those found for $B_{\text{pol}} = 0.2$, but the range of wavenumber orientations for which growth occurs decreases significantly. The pure PSS mode does not grow appreciably for small magnetic fields, but a very small-scale hydrodynamic mode becomes important for $B_{\text{pol}} < 0.05$ G; this mode has k_{ϖ}/k_Z that is similar to that of the PSS mode, and the growth rate is quite large, $\sim 0.67\Omega$. For the hydrodynamic mode, the field orientation is unimportant, because the small field means the magnetic tension does not reduce the growth rate of modes.

The strong dependence of growth rates on field orientation for PSS modes may be explained as follows: because this mode is driven by radial shear, we expect that the maximal growth rate will correspond to $k_{\varpi}/k_Z \sim -\cot\theta$. However, for fields oriented in the θ and Z directions, the values k_{ϖ}/k_Z for which $|\mathbf{k}\cdot\mathbf{v}_{\text{A}}| \ll kv_A$ are either positive or very large in magnitude, so the parameters for which shear drives modes do not coincide with parameters for which $|\mathbf{k}\cdot\mathbf{v}_{\text{A}}| \ll kv_A$. As shown in Section 4.3.2.1, this type of mode typically grows faster than the CMRI mode, so approximately radially oriented fields lead to the fastest MRI growth rates.

In Region TU, we find that no significant growth of modes with $|\mathbf{k}\cdot\mathbf{v}_{\text{A}}| \sim \Omega$ occurs for initial magnetic fields of magnitude smaller than ~ 0.002 G, but growth can occur even for large magnetic fields $B \sim 10^4$ G. The growth of modes with $|\mathbf{k}\cdot\mathbf{v}_{\text{A}}| \sim \Omega$ is inhibited by resistive and viscous dissipation for small fields, but for

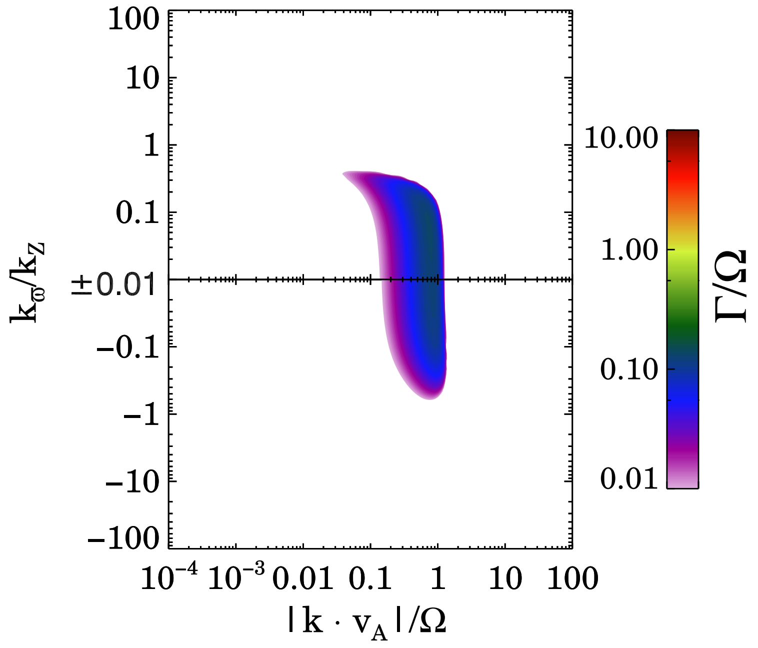

Figure 4.8 The dependence of the growth rate Γ on the phase space parameters k_ϖ/k_Z and $|\mathbf{k}\cdot\mathbf{v}_\mathrm{A}|$ at the location $r = 0.692r_\odot$, $\theta = 30^\mathrm{o}$ in Region TS for a magnetic field oriented in the θ direction. Only the CMRI mode at small $|k_\varpi/k_Z|$ is present, while there is no mode corresponding to $|\mathbf{k}\cdot\mathbf{v}_\mathrm{A}| \ll kv_A$. The fastest growth is found at $k_\varpi/k_Z = 0.068$ and $|\mathbf{k}\cdot\mathbf{v}_\mathrm{A}| = 0.753\Omega$, which corresponds to a CMRI mode.

large fields the absence of stable stratification means that modes for which $|\mathbf{k}\cdot\mathbf{v}_\mathrm{A}| \sim \Omega$ can grow even though they correspond to relatively large mode wavelengths. For these large wavelengths, the fastest-growing modes become adiabatic because their wavelengths are large enough that thermal diffusion is negligible.

The growth rate of the small-scale shear mode in Region TU at the location $r = 0.721r_\odot$, $\theta = 37.5^\mathrm{o}$ for $B_\mathrm{pol} = 0.2$ G is approximately 0.85Ω, and it varies by less than 5% with orientation. The growth rate does not change significantly for larger magnitudes of B, and the dependence on field orientation remains very small. In contrast, the growth rate decreases quickly as B is decreased, dropping to $\sim 0.01\Omega$ for $B_{\rm pol} = 0.004$ G. The dependence of growth rates on orientation is also significantly increased for small B; the growth rate for a field oriented in the Z direction is typically approximately twice that for a field oriented in the r and θ directions for $B_{\rm pol} < 0.02$G.

4.3.3 Nonaxisymmetric Effects

We now consider the effects of nonaxisymmetry on the growth rate of modes. Because all nonaxisymmetric terms in the dispersion relation are proportional to $\mathbf{k} \cdot \mathbf{v}_{\rm A}$ or $\omega_{{\rm A}\phi}$, they are typically negligible for large-scale hydrodynamic shear modes and for all types of convective modes if m is small. For very large m, the growth rate of these large-scale modes is slightly reduced by the toroidal magnetic tension, but this effect is very small. For the small-scale MRI modes in Regions TS and TU, the effects of nonaxisymmetry are considerably stronger and more complex. In general, we find that the fastest-growing small-scale shear modes are always nonaxisymmetric in these regions, although the differences in growth rate are very small for the maximum nonaxisymmetric wavenumber $|m| = 15$ used in our previous analysis.

For small-scale shear modes in regions of the Sun close to the tachocline with $\varpi \sim 0.7 r_\odot$ $\Omega \sim 2 \times 10^{-6}$ rad/s, $\rho \sim 0.2 {\rm g\ cm}^{-3}$, and the small seed fields $B \sim 1G$ and moderate ratios of toroidal to poloidal field $|R_{\rm TP}| \sim 5$ that we have used in our analysis, $\omega_{{\rm A}\phi} \sim 10^{-5}\Omega$ in the tachocline. Therefore, the explicitly nonaxisymmetric terms in Equation (4.16) are negligible, and nonaxisymmetric effects derive almost entirely from the contribution $m\omega_{A\phi}$ of the nonaxisymmetric field and wavenumber to $\mathbf{k}\cdot\mathbf{v}_A$. As a result, the strength of nonaxisymmetric effects can be parameterized using the ratio $\widetilde{\mu}$ of the toroidal and poloidal contributions to $\mathbf{k}\cdot\mathbf{v}_A$. Because $(\mathbf{k}\cdot\mathbf{v}_A)_{\rm pol} \sim \Omega$ for small-scale shear modes, we can estimate $\widetilde{\mu} \sim \mu$, where we define $\mu \equiv m\omega_{A\phi}/\Omega$. In the solar tachocline, this corresponds to $\mu \sim 10^{-5}m$. Therefore, our calculation with $|m| < 15$ corresponds to a linear regime $\mu \ll 1$ in which nonaxisymmetric effects represent a perturbation to $\mathbf{k}\cdot\mathbf{v}_A$. In this regime, the fastest-growing modes will always be nonaxisymmetric with the highest possible value for $|m|$; this is exactly what we have observed in our calculation of growth rates in Section 4.3.1.

While it is relatively easy to predict that a nonaxisymmetric mode of maximal $|m|$ will be the fastest-growing mode, it is extremely difficult to explain analytically which sign for $B_Z k_Z/(m\omega_{A\phi})$ leads to faster growth because $\mathbf{k}\cdot\mathbf{v}_A$ is present throughout the dispersion relation; we do not attempt to do so in this paper. Numerically, we find that for the CMRI mode, the sign of $B_Z k_Z/(m\omega_{A\phi})$ that reduces the value of $|\mathbf{k}\cdot\mathbf{v}_A|$ leads to the fastest growth rate. For PSS modes, the same is typically true, although the magnitude of the dependence on $B_Z k_Z/(m\omega_{A\phi})$ is typically very small. In contrast, for the mode in Region TU, the sign of $B_Z k_Z/(m\omega_{A\phi})$ that increases $|\mathbf{k}\cdot\mathbf{v}_A|$ always leads to the largest growth rate. For modes in Region TS the value of k^2 for which modes grow fastest is determined by the stable stratification for the CMRI mode, and by resistive dissipation for the PSS mode. Therefore, the reduction in magnetic tension for a given value of k may mean that the mode can grow slightly faster. For the small-scale mode in Region TU, the magnitude of $|N|$ is very small

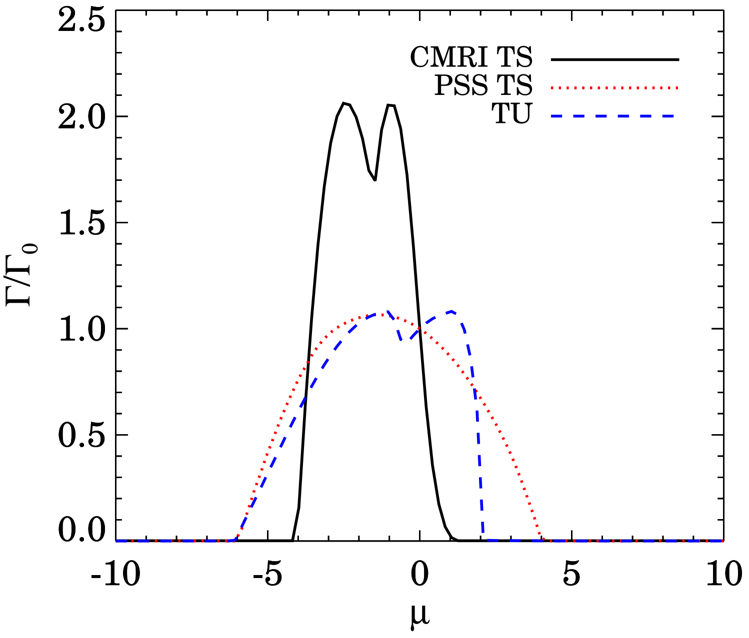

Figure 4.9 The dependence on the dimensionless toroidal wavenumber μ of the ratio of the growth rate Γ of the fastest-growing nonaxisymmetric modes with dimensionless toroidal wavenumber μ to the growth rate Γ_0 of the fastest-growing axisymmetric modes for $R_{\mathrm{TP}} = 500$. The three types of mode shown are the CMRI and PSS modes in Region TS at $r = 0.692r_\odot$, $\theta = 30^\circ$ and the small-scale mode in Region TU at $r = 0.721r_\odot$, $\theta = 37.5^\circ$.

and k is moderate in magnitude, so magnetic effects are responsible for both driving the instability and stabilizing modes via magnetic tension; on balance, the driving of modes resulting from slightly increased $|\mathbf{k} \cdot \mathbf{v}_{\mathrm{A}}|$ for a given k appears to be more important than the increase in magnetic tension.

4.3.3.1 Nonaxisymmetric Mode Growth Rates

To ascertain the effects of nonaxisymmetry for small-scale modes, we now compare the growth rate Γ of nonaxisymmetric modes with larger wavenumbers

$|m| > 15$ to the growth rate Γ_0 of the corresponding axisymmetric mode. The local approximation requires that $m/\varpi \ll k_\varpi, k_Z$, which corresponds to the condition on μ

$$\mu \ll \varpi R_{\rm TP} \min(1, |k_\varpi/k_Z|). \tag{4.56}$$

We therefore analyse nonaxisymmetric effects for toroidal-to-poloidal field ratio RTP = 500, which ensures that the local approximation is obeyed for modes with $|\mu| < 10$. Figure 4.9 shows the dependence of the growth rate Γ of the fastest-growing nonaxisymmetric modes on the normalized nonaxisymmetric wavenumber μ; the growth rate is given in terms of the growth rate Γ_0 of the fastest-growing mode with $\mu = 0$. We use the same poloidal field geometry as in our axisymmetric analysis, with $B_{\rm pol} = 0.2$G and $B_\varpi/B_Z = \tan\theta$, which corresponds to a magnetic field oriented in the r direction, and set the sign $B_Z k_Z/(\omega_{{\rm A}\phi}) > 0$. For $\mu \ll 1$, the size of nonaxisymmetric effects is small and the variation with μ is approximately linear. The maximum value of Γ corresponds to $\mu < 0$ for both modes with $r = 0.692 r_\odot$, $\theta = 30^{\rm o}$ in Region TS, and $\mu > 0$ for the mode in Region TU at $r = 0.721 r_\odot$, $\theta = 37.5^{\rm o}$. Both of these results are in agreement with our general discussion of the linear regime in Section 4.3.3.

For $|\mu| > 1$, nonaxisymmetric effects become nonlinear, and the growth rate of nonaxisymmetric modes peaks and then goes down as the tension produced by the toroidal field begins to stabilize the mode. For all modes, this stabilization occurs at smaller $|\mu|$ for $\mu < 0$ than for $\mu > 0$; this is because we have used the sign convention $(\mathbf{k} \cdot \mathbf{v}_{\rm A})_{\rm pol} > 0$, so that large negative μ corresponds to a smaller value for $|\mathbf{k} \cdot \mathbf{v}_{\rm A}|$ than does large positive μ. As shown in Figure 4.9, the maximum growth

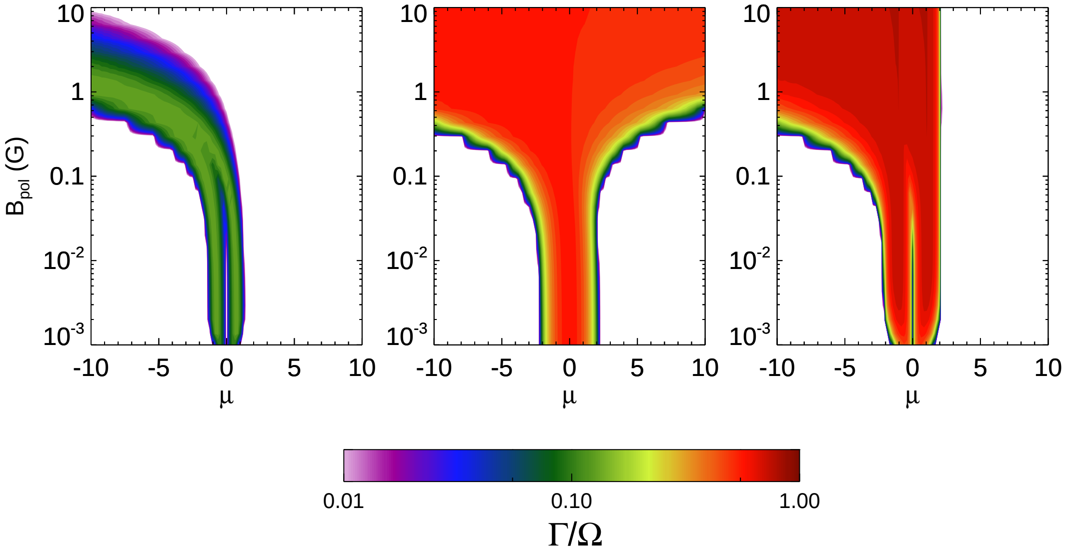

Figure 4.10 The dependence of the growth rate Γ of MRI modes on the poloidal magnetic field B_{pol} and the dimensionless toroidal wavenumber μ for the CMRI (left panel) and PSS mode (middle panel) in Region TS at $r = 0.692 r_{\odot}$, $\theta = 30^{\circ}$ and the small-scale mode in Region TU (right panel) at $r = 0.721 r_{\odot}$, $\theta = 37.5^{\circ}$. The increase in the range of μ with B_{pol} reflects the fact that because R_{TP} is held constant, $\mu \propto B_{\mathrm{pol}}$.

rate of nonaxisymmetric modes is $2.07\Gamma_0$ for the CMRI mode, $1.09\Gamma_0$ for the mode in Region TU, and $1.07\Gamma_0$ for the PSS mode. For the CMRI mode and the mode in Region TU, two peaks are present in the growth rate, corresponding to values of μ that produce values of $\mathbf{k} \cdot \mathbf{v}_{\mathrm{A}}$ of the same magnitude but opposite sign. The locations of these peaks are offset from $\mu = 0$ in the direction of increasing growth rate in the linear regime. For these modes, nonaxisymmetric effects thus correspond to an adjustment in $\mathbf{k} \cdot \mathbf{v}_{\mathrm{A}}$ that breaks its degeneracy with the magnitude of k, and the peaks correspond to the values of $|\mathbf{k} \cdot \mathbf{v}_{\mathrm{A}}|$ that are optimized for fast growth, rather than being determined by dissipative constraints on k. The dependence of growth rate on μ is significantly weaker for the PSS mode than for the other two modes, and only one peak in growth rate is present. This is a result of the strong dependence of the growth of PSS modes on k_ϖ/k_Z, which is required to ensure that the large magnitude of k does not stabilize the mode via magnetic tension. When a nonaxisymmetric wavenumber $\mu \neq 0$ is introduced, its contribution to $\mathbf{k} \cdot \mathbf{v}_\mathrm{A}$ can be canceled by shifting the k_ϖ/k_Z slightly; for $r = 0.692 r_\odot$, $\theta = 30^\mathrm{o}$, the value of k_ϖ/k_Z is shifted from the axisymmetric value of -1.57 to -1.20 for the fastest-growing modes at large $|\mu|$. Thus, μ is degenerate with the parameter k_ϖ/k_Z for the PSS modes, and the degeneracy between k and $\mathbf{k} \cdot \mathbf{v}_\mathrm{A}$ is not broken by nonaxisymmetric effects, which are therefore reduced in importance.

4.3.3.2 Effects of Field Geometry on Nonaxisymmetric Modes

The field geometry in the nonaxisymmetric case is determined by three parameters: the direction of the poloidal field B_ϖ/B_Z, the toroidal-to-poloidal field ratio R_TP, and the magnitude of the poloidal field B_pol. We find that the effects of B_ϖ/B_Z on the nonaxisymmetric small-scale shear modes are similar to those on their axisymmetric counterparts; the PSS mode exists only for mode directions that are close to radial, while growth rates for the CMRI mode and the small-scale shear mode in Region TU vary by factor of 2 at most with direction. The initial toroidal-to-poloidal field ratio R_TP has been subsumed into the dimensionless wavenumber μ; once this has been done, the only effect R_TP then has is in determining the range of μ for which our linear analysis is valid. The poloidal field magnitude B_pol is therefore the only field geometry parameter that has significant implications for the strength of nonaxisymmetric effects in our linear analysis.

Figure 4.10 shows the dependence of the growth rate of the three small-scale shear modes on the poloidal magnetic field B_{pol} and the dimensionless toroidal wavenumber μ for toroidal-to-poloidal field ratio $R_{\mathrm{TP}} = 500$. The typical dependence on μ for all modes is similar to that described in the previous section, including a double peak in growth rate for the CMRI mode and the mode in Region TU, and a single peak for the PSS mode for most values of B_{pol}. For the CMRI mode, which is shown in the left panel of Figure 4.10, the nonaxisymmetric growth rate is larger than the axisymmetric growth rate by a very large factor of > 100 for small initial poloidal field $B_{\mathrm{pol}} \sim 0.001$. As B_{pol} is increased, the ratio of the growth rates of the fastest-growing nonaxisymmetric and axisymmetric modes eventually becomes smaller, reaching $\Gamma/\Gamma_0 \sim 2$ for $B_{\mathrm{pol}} \sim 1$G. As B_{pol} is increased further, the poloidal magnetic tension begins to stabilize the growth of modes, as described in Section 4.3.2.5. For $B_{\mathrm{pol}} > 1$G, axisymmetric modes are stabilized completely, but modes with $\mu < 0$ are able to grow because the nonaxisymmetric contribution to $\mathbf{k} \cdot \mathbf{v}_{\mathrm{A}}$ reduces the impact of the poloidal tension. For the mode in Region TU, shown in the right panel of Figure 4.10, the nonaxisymmetric growth rate is again larger than the axisymmetric growth rate by a very large factor of > 100 for small initial poloidal field $B_{\mathrm{pol}} \sim 0.001$. As B_{pol} is increased, the ratio Γ/Γ_0 of the growth rates of the fastest-growing nonaxisymmetric and axisymmetric modes decreases quickly, falling below 1.2 for poloidal fields $B_{\mathrm{pol}} \sim 0.2$G; however, nonaxisymmetric modes remain the fastest-growing modes. For even larger B_{pol}, fast nonaxisymmetric growth occurs for large negative μ, but not for large positive μ; this is because negative μ reduces the magnetic tension.

The typical evolution of the PSS mode is very different from that of the other two modes. For all values of B_{pol}, the nonaxisymmetric and axisymmetric modes have very similar growth rates. For $B_{\mathrm{pol}} < 10^{-2}$ the observed PSS mode growth is that of the hydrodynamic mode discussed in Section 4.3.2.5, and axisymmetric modes grow faster than nonaxisymmetric modes; for larger B_{pol}, the observed growth corresponds to the PSS mode proper, and nonaxisymmetric modes grow slightly faster than axisymmetric modes. The weak dependence of growth rate on μ for all B_{pol} reflects the degeneracy between μ and k_{ϖ}/k_Z for the PSS mode.

4.4 Discussion

We now compare our results to other authors and discuss the nonlinear effects leading to the saturation of the magnetic field.

4.4.1 Comparison with Other Studies

Our axisymmetric study of modes in the Sun is most directly comparable to the research of Parfrey & Menou (2007) and Masada (2011). Parfrey & Menou (2007) studied the growth of modes in the stably stratified tachocline, and found that significant growth of instability occurred for $\theta < 53^{\mathrm{o}}$, and that initially radial fields lead to faster growth than do toroidal fields over most of the domain. The trends we find are similar; however, by studying the phase space structure of the dispersion relation we have identified the two small-scale shear modes in Region TS and found that the reason radial fields lead to faster growth is that they allow the fast growth of PSS modes. Quantitatively, the growth rates found by Parfrey & Menou (2007) are somewhat larger than ours in Region TS, because we use a larger value for N corresponding to the bottom of the tachocline, while they use a smaller value corresponding to its center.

Masada (2011) investigated the growth of axisymmetric modes throughout the Sun in the absence of thermal buoyancy; thus, the only unstable modes present were MRI modes. He found that the MRI is unstable only near the tachocline at high latitudes and very close to the surface. The stabilizing effect of density stratification was substantially reduced by doubly–diffusive effects in the convectively stable portion of the tachocline. He found that the layer near the surface is formally unstable to the MRI, but argued that the dispersion relation that is the basis of the instability criteria should not be applied in a region of vigorous convection since the stationary background assumed to derive the dispersion relation does not exist there. We have confirmed using a linear analysis of the full dispersion relation that modes driven by shear are not present in the outer parts of the convection zone, which corresponds to Region C. This is because for moderate seed fields $B \sim 1$ G, modes with $\mathbf{k} \cdot \mathbf{v}_{\mathrm{A}} \sim \Omega$ that correspond to the shear modes also have $k^2 \xi < N$; therefore, hydrodynamic convective modes with these parameters have a growth rate $\sim |N|$, and there are no parameters for which shear modes are significant.

Our nonaxisymmetric results can be most directly compared to those found by Masada et al. (2007) in the context of stably stratified proto-neutron stars. They found that the fastest growth of nonaxisymmetric modes occurred for $\mu \sim 1$, and that the nonaxisymmetric modes grew much faster than axisymmetric modes unless the poloidal field was very large. We find that the relative growth rates of nonaxisymmetric and axisymmetric modes depend on the type of mode. For the CMRI mode, the nonaxisymmetric modes always grow significantly faster than axisymmetric modes, especially for very large and very small poloidal fields. For the PSS mode, nonaxisymmetric modes never grow more than 5% faster than axisymmetric modes; this mode is not detected by Masada et al. (2007), because their dispersion relation assumes that $B_{\varpi} = 0$. Finally, for the mode in the convectively unstable Region TU, nonaxisymmetric modes grow much faster than axisymmetric modes only for very small poloidal fields. Masada et al. (2007) do not investigate the CMRI case with very large fields, for which nonaxisymmetric CMRI modes are dominant because the axisymmetric modes are stabilized by magnetic tension; therefore, our results are consistent within their range of validity.

4.4.2 Modes in Region TL and the Emergence of Active Regions

Active regions on the Sun typically emerge near a central latitude which varies from $\sim 30^{\circ}$ at the beginning of a solar cycle to 0° at its end. The typical spread in the central latitude of emergence for large active regions is $\pm 10^{\circ}$, while our results in Region TL suggest that the growth of modes is significantly suppressed at latitudes smaller than 7°. If shear modes corresponding to the MRI also suppress the formation of large-scale features at upper latitudes corresponding to $\theta \leq 53^{\circ}$, as suggested by Parfrey & Menou (2007), we expect that features will be produced close to the tachocline primarily at latitudes in the range 7°–35°. If large-scale dynamo effects restrict the emergence of these structures on the surface to within

$\pm 10^{\mathrm{o}}$ of the central emergence latitude, then we expect that more active regions will emerge when the central emergence latitude is close to the middle of its range, 15^{o}, than when it is at extreme latitudes of 0^{o} or 30^{o}. This is consistent with observations: solar maximum occurs when the central emergence latitude is $\sim 15^{\mathrm{o}}$, while solar minimum occurs when the central emergence latitude is close to 0^{o} or 30^{o}.

4.4.3 MRI Saturation

When the MRI is initiated, the field may be small, $\omega_A \ll q\Omega$, but that condition will not last long as the field grows exponentially. The field may be susceptible to tearing by associated thermal convection, but convection will be ineffective on large scales, in excess of the pressure scale height. Convective timescales may be shorter than rotational timescales, as they are over much of the solar convective zone, but only on the scale of a convective eddy. The large-scale structure of the solar magnetic field will depend on the interaction of local convective and MRI eddies with global dynamo effects, so the determination of the saturated field is nontrivial. Nevertheless, it is useful to consider the saturated field produced by these individual effects separately.

In the absence of thermal convection, saturation of the MRI occurs for $v_{\mathrm{A}} \sim q\Omega r$ or $\omega_{\mathrm{A}} \sim q\Omega$ (Balbus & Hawley, 1998; Vishniac, 2009). The associated saturation field $B_{\mathrm{sat,MRI}}$ is then

$$B_{\mathrm{sat,MRI}} \sim \sqrt{4\pi\rho}|q|\Omega\varpi. \tag{4.57}$$

That condition will be reached quickly, on the timescale Ω^{-1}. While linear field winding that is responsible for the Ω effect in the dynamo context might be active in the tachocline, the MRI grows exponentially, and may thus be responsible for the level of fields that are thought to then be driven to the solar surface by buoyancy. If so, the field winding would have no effect on the growth of field; it would only be responsible for creating large-scale toroidal fields from the strong small-scale poloidal fields produced by the MRI. For the angular velocity of the tachocline, $\Omega_{\text{tach}} = 2.7 \times 10^{-6}$ rad s^{-1}, and a density $\rho_{\text{tach}} = 0.2$ g cm^{-3}at a spherical radius of $r_{\text{tach}} = 5 \times 10^{10}$ cm, the saturation value of the MRI is of order $B_{\text{sat}} \sim 2.4 \times 10^5 |q|$ G. Since q is of order unity in the tachocline, if the magnetic field at $\theta < 53^{\text{o}}$ near the solar surface originates in the tachocline, it could easily be produced by the MRI.

In the absence of shear, the saturation field will be in equipartition with the turbulent pressure resulting from convection, so that $v_{\text{A}} \sim v_{\text{conv}}$. Thus, the saturated magnetic field $B_{\text{sat,conv}}$ resulting from convection will be

$$B_{\text{sat,conv}} \sim \sqrt{4\pi\rho} v_{\text{conv}}. \tag{4.58}$$

The convective velocities in the solar envelope were previously thought to range from about 0.05 km s^{-1} at the base of the convective zone to about 2.5 km s^{-1} at the top of the envelope (Christensen-Dalsgaard et al., 1996; Howe et al., 2005). Recent studies by Hanasoge et al. (2010) and Hanasoge et al. (2012) have shown, however, that on large scales, the convective velocities are typically smaller than 0.01 km s^{-1} near the solar surface at $r > 0.96 r_\odot$. The saturated field from large-scale convection at $r = 0.96 r_\odot$, where $\rho \sim 0.01$ g cm^{-3}, is thus only $B_{\text{sat,conv}} \sim 3.5 \times 10^2$ G. Thus, the saturation field produced by the MRI at high latitudes as the tachocline merges into the fully convective envelope would easily dominate any field that was in equilibrium with the turbulent pressure. Even if small-scale convective velocities near the solar surface are ~ 1 km s^{-1} as suggested by the earlier studies, the resulting field at $r = 0.96r_\odot$ is $B_{\rm sat,conv} \sim 3.5 \times 10^4$ G, which is still somewhat smaller than that produced by the MRI in the tachocline.

Because convection on large scales is so weak near the solar surface, MRI modes in Region C may be of greater importance than suggested by our linear analysis. In particular, the region of unstable MRI growth near the solar surface at lower latitudes noted by Masada (2011) may actually play an important role in the origin of the solar magnetic field. Because $q \sim 1$ in this shear layer, the predicted magnetic field from the growth of MRI modes in situ at $r = 0.96r_\odot$ is approximately $B_{\rm sat,MRI} \sim 10^4 G$. This field is significantly larger than the field produced by convection according to recent large-scale helioseismology results, and comparable to the field produced by convection according to the earlier studies. Thus, the MRI may play an important role in the origin of small-scale magnetic fields at all latitudes in the Sun.

4.5 Conclusions

In this paper, we have derived a dispersion relation for nonaxisymmetric instability including the MRI and used it to calculate the growth rate of modes throughout the Sun. We have explored the phase space defined by the magnitude and direction of the wavenumber **k**, identifying the fastest-growing modes at each location in the Sun, and explored the dependence of these growth rates on the initial magnitude and direction of the magnetic field **B**. We find that nonaxisymmetric effects typically represent a perturbation to the axisymmetric modes for toroidal-to-poloidal field ratios of ~ 5 that are typical in stellar field equilibria, so we first analyse the axisymmetric modes and then explore how the resulting conclusions are changed by nonaxisymmetric effects. Unless otherwise stated, specific numerical values in this section are based on an initial magnetic field with poloidal components oriented in the r direction, toroidal-to-poloidal field ratio $R_{\mathrm{TP}} = 5$, and poloidal field magnitude $B_{\mathrm{pol}} = 0.2\mathrm{G}$. Our conclusions are as follows:

- The overall instability contains three types of submodes: hydrodynamic convective modes, hydrodynamic shear modes, and small-scale MRI modes. The hydrodynamic modes are large-scale with wavelengths on the order of the pressure scale height, while the magnetohydrodynamic modes, which can be called MRI modes because they have $\mathbf{k} \cdot \mathbf{v}_{\mathrm{A}} \sim \Omega$ and are driven by shear, grow on much smaller length scales. The typical growth rates of the convective modes are on the order of the Brunt–Väisälä frequency N, while the typical growth rates of the shear modes are typically $\lesssim \Omega$.
- Those parts of the Sun in which significant growth of modes occurs may be divided into four regions in which the properties of these modes are significantly different: Region TS, which is located in the stably stratified tachocline, Region TU, which is located in the convectively unstable tachocline and lower convection zone at colatitudes $0^{\mathrm{o}} < \theta < 60^{\mathrm{o}}$, Region TL, which is located in the convectively unstable tachocline and lower convection zone at colatitudes $\theta > 60^{\mathrm{o}}$, and Region C, which is located in the upper convection zone at all latitudes.

- The typical growth rates of the fastest-growing modes are $\sim \Omega$ in Regions TS and at smaller radii in Region TU, $\sim 0.1\Omega$ in Region TL, and $\sim N$ in Region C. These modes are driven by shear in Region TS and at the bottom of Region TU in the tachocline, by both shear and convection in the upper parts of Region TU and at the bottom of Region TL, and solely by convection in Region C and the upper parts of Region TL. In general, the importance of shear is greatest in the tachocline and typically decreases outward, while the strength of convection increases as one moves outward from the tachocline.

- In Region TS, all modes that grow significantly are small-scale MRI modes with $|\mathbf{k} \cdot \mathbf{v}_\mathrm{A}| \sim \Omega$ so that thermal diffusion can reduce the stabilization resulting from stable stratification, and all of these modes are driven by shear. Two types of MRI modes exist in this region: the CMRI mode with $|k_\varpi/k_Z| < 1$ and typical growth rate $\sim 0.1\Omega$ and the PSS mode with large k, poloidal wavenumber nearly perpendicular to the magnetic field, and typical growth rate $\sim \Omega$.

- In Region C, the growth rate of modes is $\sim |N|$ for all $\mathbf{k}$ such that $k^2\xi < |N|$. The growth rate does not depend strongly on k_ϖ/k_Z except at two locations: for $k_\varpi/k_Z \approx \tan\theta$, the growth of modes is entirely suppressed for $k_\varpi/k_Z \approx \tan\theta$ and is slightly enhanced $k_\varpi/k_Z \approx -\cot\theta$. For $k_\varpi/k_Z \approx \tan\theta$, the perturbed velocity is perpendicular to gravity and no growth occurs, while for $k_\varpi/k_Z \approx -\cot\theta$ the perturbed velocity is parallel to gravity, and the growth of instability is maximized.

- In Region TL, hydrodynamic shear and convective modes exist on large scales for $|k_\varpi/k_Z| \gg 1$, so these modes drive flows in the Z direction at nearly constant ϖ. Because the radial shear is positive in Region TL, modes with $|k_\varpi/k_Z| < \tan\theta$ that cause flows to move in the ϖ direction are stabilized by an angular momentum barrier. As a result, the growth rates of the fastest-growing modes in Region TL are much smaller than those in any other region, typically smaller than 0.1Ω. The suppression of large-scale modes near the equator, combined with the suppression of large-scale modes by small-scale shear modes in Regions TS and TU suggested by Parfrey & Menou (2007), may explain why active regions tend to appear at latitudes of $\approx 20^{\mathrm{o}}$ in the Sun during solar maximum.

- In Region TU, we observe four modes, two of which are driven by convection and two of which are driven by shear. The two types of convective modes are a large-scale hydrodynamic mode of the type observed in Region TL and a weak, highly overstable mode with $k^2\xi \sim N$. The two types of shear mode are a large-scale hydrodynamic mode similar to that observed in Region TL, and a small-scale shear mode with similar properties to that of the CMRI mode. In the upper tachocline, the shear modes grow much more quickly than the convective modes, and the fastest-growing mode is the hydrodynamic shear mode. In contrast, in the lower convection zone, the fastest-growing modes are large-scale hydrodynamic modes driven by both shear and convection, with convection being slightly more important.

- The growth rate of the small-scale MRI modes has a significant dependence on the initial magnetic field strength and orientation. For the CMRI mode in Region TS, the magnetic field must be large enough that resistive and viscous dissipation is unimportant on scales for which $k^2\xi > N$ but small enough that magnetic tension does not stabilize the growth of modes on these small scales; Therefore, at the bottom of Region TS, the growth of field occurs quickly only for a relatively small range of poloidal fields: 0.02 G$< B_{\rm pol} < 0.6$ G. For the PSS mode, the only restriction on the magnetic field is that $B > 0.08$G; for smaller fields, a hydrodynamic mode is present that does not depend on field orientation. Finally, for the small-scale mode in Region TU the lack of stratification means that there is no upper limit to the magnetic field, as modes can grow on large scales; typically modes can grow for all $B_{\rm pol} > 0.1$G. Only the PSS modes in Region TS have a strong dependence on the field orientation: they only exist for fields that are within about $\pm 45^{\rm o}$ of an orientation that is close to radial, so that $\mathbf{k}\perp\mathbf{B}$ also corresponds to $\mathbf{k}$ that drives flows in the direction of the local shear, which is approximately the r direction.

- Nonaxisymmetric small-scale MRI modes with $m \neq 0$ grow faster than their axisymmetric counterparts in Regions TS and TU. For the solar case, the explicitly nonaxisymmetric terms are typically unimportant, and the only important nonaxisymmetric effect is the influence of the normalized toroidal wavenumber μ on the value of $\mathbf{k}\cdot\mathbf{v}_{\rm A}$. The primary effect of nonaxisymmetry for CMRI modes and modes in Region TU is to break the degeneracy between $\mathbf{k}\cdot\mathbf{v}_{\rm A}$ and k, substantially increasing the growth rate for CMRI modes; nonaxisymmetric effects are significantly weaker for the PSS modes, because changes in μ can be balanced by adjustments in k_{ϖ}/k_Z.

- The poloidal field magnitude $B_{\rm pol}$ has important effects on the relative growth rate of axisymmetric and nonaxisymmetric MRI modes. For the CMRI mode, nonaxisymmetric modes always have a growth rate at least twice that of the axisymmetric modes, but for very large and very small $B_{\rm pol}$, nonaxisymmetric modes are larger by a factor of > 100. For the PSS mode, nonaxisymmetric modes always have similar growth rates to axisymmetric growth modes. For the small-scale shear mode in Region TU, nonaxisymmetric growth rates are similar to axisymmetric growth rates except for small $B_{\rm pol}$, where they can be larger by a factor of > 100.

In this paper, we have applied a local WKB analysis to calculate the growth rates of modes. This technique has two important limitations that can be remedied in future work. The first limitation of this analysis is that the constraints on the value of μ implied by the local approximation have made it impossible to explore the full parameter space of nonaxisymmetric modes for small $R_{\rm TP} \sim 5$. An eigenvalue analysis of the full MRI would be an important next step that would make it possible to determine whether the dependence of growth rate on μ is the same for all toroidal fields.

The second and more serious shortcoming of the local WKB analysis is that it cannot explore the nonlinear effects of MRI modes. The manner in which the physics of the MRI would affect the solar dynamo and observed field effects at the surface is a complex problem. From our analysis it seems that the physics of the

MRI should be considered in this context in both the tachocline and in the lower parts of the convective envelope itself.

There is a general perception that the MRI is less important in stars than the Tayler-Spruit mechanism (Pitts & Tayler, 1985; Spruit, 1999, 2002) because the threshold of required shear for the latter in the radiative zone is lower than for the MRI. While the MRI may be unimportant if shear grows very slowly, it is uncertain that this will be the case for all stars; if strong shear can be produced in a stellar model, the presence of diffusion allows the MRI to grow quickly in radiative regions. Direct comparison between the two instabilities shows that the MRI grows exponentially rapidly on the timescale Ω^{-1} in any environment that has strong shear, while the typical timescale for the growth of the Tayler-Spruit instability is $\sim \Omega\omega_{\mathrm{A}\phi}^{-2}$ for weak fields with $\omega_{\mathrm{A}\phi} \ll \Omega$ (Pitts & Tayler, 1985). Thus, if some environments in stars are unstable to the MRI, the growth rate for the MRI is likely to be much faster than that of the Tayler-Spruit instability, by a factor of $(\Omega/\omega_{\mathrm{A}\phi})^2$. Parfrey & Menou (2007) have previously noted that the tachocline is probably also unstable to the Tayler-Spruit mechanism, but argue that if both instabilities are present, the MRI is likely to dominate because of its much faster growth rate. It is also commonly assumed that instabilities that depend on shear will not grow in convective zones (e.g., Heger et al., 2005). MRI modes in convective zones are not hindered by stable stratification and can grow even when only moderate shear is present; because large-scale convective motions may be slow in convective regions, it is uncertain that the presence of convection will prevent MRI modes from growing, either.

These considerations suggest that the role of the MRI in the evolution of stars in general deserves more attention. The saturation fields resulting from the MRI can be appreciable, and this may call for magnetic buoyancy effects to be re-evaluated. The MRI, as well as the Tayler-Spruit mechanism and other dynamo processes, may also leave behind fossil fields in one stage of evolution that affect the physical conditions at later stages of evolution. The role of magnetic fields in stellar evolution remains a major challenge requiring fully three-dimensional studies. This paper, investigating the combined effects of shear and convection as well as nonaxisymmetric effects in the Sun, may constitute a guide in constructing these studies.

Chapter 5

Future Research

In this dissertation, we have investigated three instabilities of magnetic fields in astrophysical systems and the resulting energy conversion. Several future projects are feasible, guided by the research in this dissertation. One of these projects focuses on studying how reconnection is initiated, guided by observations. In particular, reconnection in a magnetized jet may be understood as occurring between randomly oriented, finite, turbulent magnetic field domains on various length scales. One possible avenue of research involves modeling this turbulent reconnection by modifying models of multiscale reconnection to take account of the three dimensional isotropy of current sheets that is observed in the particle simulations in Chapter 3 of this dissertation. Calculations of the expected particle energy spectrum for such a multiscale reconnection hierarchy could then be compared with observed GRB emission. In contrast to this picture, observations of solar flares indicate that the most common sites of reconnection in the solar corona are associated with separators that connect pairs of null points with zero magnetic field. One possible project uses a separator configuration to initialize PIC simulations that can determine the structure of reconnection regions and the efficiency and mechanisms of particle acceleration in this realistic case. A first step in this direction has already been taken by Baumann et al. (2013), who carried out particle-in-cell simulations of reconnection and particle acceleration in a solar null point configuration. Continued research would make it possible to gain a deeper understanding of observed solar eruptions and the high-energy particles produced by those eruptions.

A second possible area of research stimulated by this dissertation involves exploring the effect of the magnetorotational instability (MRI) on stellar evolution. One possible project would involve developing instability criteria and an effective viscosity prescription for the effects of the MRI on the evolution of stars in both convective and radiative regions, guided by the observed MRI growth rates in the present Sun presented in Chapter 4 of this dissertation. One could then use a one-dimensional stellar evolution code such as MESA to model the evolution of high-mass stars with time, including the effects of the MRI. New pre-supernova models that take into account these effects will make it possible to predict the rotation rates of the resulting collapsed cores, which will constrain the properties of the supernova and allow the determination of whether a gamma-ray burst is likely to be produced by this evolution. We hope that this dissertation will stimulate research that will lead to a greater understanding of the effects of magnetic fields in astrophysics.

Bibliography

Acheson, D. J. 1978, Royal Society of London Philosophical Transactions Series A, 289, 459

Aly J. J., 1991, Astrophys. J., 375, L61

Baalrud, S. D., Bhattacharjee, A., & Huang, Y.-M. 2012, Phys. Plasmas, 19, 022101

Balbus, S. A. 1995, Astrophys. J., 453, 380

Balbus, S. A. & Hawley, J. F. 1991, Astrophys. J., 376, 214

Balbus, S. A., & Hawley, J. F. 1994, Mon. Not. Roy. Astr. Soc., 266, 769

Balbus, S. A. & Hawley, J. F. 1998, Rev. Mod. Phys., 70, 1

Balogh, A., & Erdõs, G. 2013, Space Sci. Rev., 176, 177

Basu, S., & Antia, H. M. 2001, Mon. Not. Roy. Astr. Soc., 324, 498

Baumann, G., Haugbølle, T., & Nordlund, A. A. 2013, Astrophys. J., 771, 93

Bell, A. R. 1978, Mon. Not. Roy. Astr. Soc., 182, 147

Berger, M. A. 1997, J. Geophys. Res., 102, 2637

Berger, M. A. 1999, Plas. Phys. Cont. Fus., 41, B167

Bessho, N. & Bhattacharjee, A. 2005, Phys. Rev. Lett., 95, 245001

Bessho, N. & Bhattacharjee, A. 2007, Phys. Plasmas, 14, 056503

Bessho, N. & Bhattacharjee, A. 2010, Phys. Plasmas, 17, 102104

Bessho, N. & Bhattacharjee, A.. 2012, Astrophys. J., 750, 129

Bhattacharjee A., 2004, ARA&A, 42, 365

Blandford, R. D. & Ostriker, J. P. 1978, Astrophys. J., 221, L29

Braithwaite, J. 2009, Mon. Not. Roy. Astr. Soc., 397, 763

Brown, B. P., Miesch, M. S., Browning, M. K., Brun, A. S., & Toomre, J. 2011, Astrophys. J., 731, 69

Cerutti, B., Uzdensky, D. A., & Begelman, M. C. 2012a, Astrophys. J., 746, 148

Cerutti, B., Werner, G. R., Uzdensky, D. A., & Begelman, M. C. 2012b, Astrophys. J., 754, L33

Chanmugam, G. 1979, Mon. Not. Roy. Astr. Soc., 187, 769

Chandrasekhar, S. 1960, Proc. Nat. Acad. Sci., 46, 253

Che, H., Drake, J. F., & Swisdak, M. 2011, Nature, 474, 184

Chen J., 2001, Space Sci. Rev., 95, 165

Chen, L., Bhattacharjee, A., Puhl-Quinn, P. A., et al. 2008, Nature Phys., 4, 19

Christensen-Dalsgaard, J., et al. 1996, Science, 272, 1286

Clyne, J., Mininni, P., Norton, A., & Rast, M. 2007, New J. Phys., 9, 301

Dasso S., Mandrini C. H., Démoulin P., Luoni M. L., Gulisano A. M., 2005, Adv. Space Res., 35, 711 Dennis B. R., Haga L., Medlin D. A., Tolbert A. K., 2006, Bull. Amer. Astr. Soc., 38, 2

Daughton, W. 1999, Phys. Plasmas, 6, 1329

Daughton, W. & Karimabadi, H. 2007, Phys. Plasmas, 14, 072303

Daughton, W., Roytershteyn, V., Karimabadi, H., et al. 2011, Nature Phys., 7, 539

Dikpati, M., & Gilman, P. A. 2001, Astrophys. J., 551, 536

Drake, J. F., Opher, M., Swisdak, M., & Chamoun, J. N. 2010, Astrophys. J., 709, 963

Drake, J. F., Swisdak, M., Che, H., & Shay, M. A. 2006, Nature, 443, 553

Egedal, J., Daughton, W., Drake, J. F., Katz, N., & Lê, A. 2009, Phys. Plasmas, 16, 050701

Egedal, J., Daughton, W., & Le, A. 2012, Nature Phys., 8, 321

Emslie A. G., Dennis B. R., Holman G. D., Hudson H. S., 2005, J. Geophys. Res. - Space Phys., 110, 11103

Endal, A. S., & Sofia, S. 1981, Astrophys. J., 243, 625

Fermo, R. L., Drake, J. F., & Swisdak, M. 2010, Phys. Plasmas, 17, 010702

Fabian, A. C. 2012, Ann. Rev. Ast. Astrophys., 50, 455

Forbes T. G., Linker J. A., Chen J., et al., 2006, Space Sci. Rev., 123, 251

Fredvik T., Kjeldseth-Moe O., Haugan S. V. H., Brekke P., Gurman J. B., Wilhelm K., 2002, Adv. Space Res., 30, 635

Fricke, K. 1969, Astrophys. Lett., 3, 219

Fu, X. R., Lu, Q. M., & Wang, S. 2006, Phys. Plasmas, 13, 012309

Galeev A. A., Rosner R., Vaiana G. S., 1979, Astrophys. J., 229, 318

Giannios, D. 2010, Mon. Not. Roy. Astr. Soc., 408, L46

Goldreich, P., & Schubert, G. 1967, Astrophys. J., 150, 571

Gopalswamy N., Mikić Z., Maia D., et al., 2006, Space Sci. Rev., 123, 303

Hanasoge, S. M., Duvall, T. L., Jr., & DeRosa, M. L. 2010, Astrophys. J. , 712, L98

Hanasoge, S. M., Duvall, T. L., & Sreenivasan, K. R. 2012, Proc. Nat. Acad. Sci., 109, 1192

Heger, A., Woosley, S. E., & Spruit, H. C. 2005, Astrophys. J., 626, 350

Hesse, M., Neukirch, T., Schindler, K., Kuznetsova, M., & Zenitani, S. 2011, Space Sci. Rev., 160, 1–4, 3.

Hesse, M. & Zenitani, S. 2007, Phys. Plasmas, 14, 112102

Hoshino, M. 2012, Phys. Rev. Lett., 108, 135003

Hoshino, M., Mukai, T., Terasawa, T., & Shinohara, I. 2001, J. Geophys. Res., 106, 25979

Howe, R. 2009, Living Reviews in Solar Physics, 6, 1

Howe, R., Christensen-Dalsgaard, J., Hill, F., Komm, R., Schou, J., & Thompson, M. J. 2005, Astrophys. J., 634, 1405

Huang, C., Lu, Q., & Wang, S. 2010, Phys. Plasmas, 17, 072306

Jaroschek, C. H., Treumann, R. A., Lesch, H., & Scholer, M. 2004, Phys. Plasmas, 11, 1151

Kim, W.-T., & Ostriker, E. C. 2000, Astrophys. J., 540, 372

Kjeldseth-Moe O., Brekke P., 1998, Sol. Phys., 182, 73

Kowal, G., de Gouveia Dal Pino, E. M., & Lazarian, A. 2011, Astrophys. J., 735, 102

Krall, N. A. & Trivelpiece, A. W. 1973, Principles of Plasma Physics (McGraw-Hill)

Krall J., Chen J., Santoro R., 2000, Astrophys. J., 539, 964

Larrabee, D. A., Lovelace, R. V. E., & Romanova, M. M. 2003, Astrophys. J., 586, 72

Liu, W., Li, H., Yin, L., Albright, B. J., Bowers, K. J., & Liang, E. P. 2011, Phys. Plasmas, 18, 052105

Lyubarsky, Y. & Kirk, J. G. 2001, Astrophys. J., 547, 437

Lyubarsky, Y. & Liverts, M. 2008, Astrophys. J., 682, 1436

Lyutikov, M. 2003, Mon. Not. Roy. Astr. Soc., 346, 540

Maeder, A. 2009, Physics, Formation and Evolution of Rotating Stars (Springer Berlin Heidelberg)

Maeder, A., & Meynet, G. 2005, Astron. & Astrophys., 440, 1041

Mahajan S. M., Miklaszewski R., Nikol'Skaya K. I., Shatashvili N. L., 2001, Phys. Plasmas, 8, 1340

Mahajan S. M., Shatashvili N. L., Mikeladze S. V., Sigua K. I., 2005, Astrophys. J., 634, 419

Mahajan S. M., Yoshida Z., 1998, Phys. Rev. Lett., 81, 4863

Markidis, S., Henri, P., Lapenta, G., et al. 2012, Nonlin. Proc. Geophys., 19, 145

Masada, Y. 2011, Mon. Not. Roy. Astr. Soc., 411, L26

Masada, Y., Sano, T., Shibata, K. 2007, Astrophys. J., 655, 447

Masada, Y., Sano, T., Takabe, H. 2006, Astrophys. J., 641, 447

McKee, C. F., & Ostriker, E. C. 2007, Ann. Rev. Ast. Astrophys., 45, 565

McKinney, J. C. & Uzdensky, D. A. 2012, Mon. Not. Roy. Astr. Soc., 419, 573

Menou, K., Balbus, S. A., & Spruit, H. C. 2004, Astrophys. J., 607, 564

Mewaldt R. A., et al., Cohen, M. S., Mason, G. M. et al. 2005, in Acharya,B. S., Gupta S., Jagadeesan P., Jain A., et al., eds, 29th International Cosmic Ray Conference. Tata Institute of Theoretical Astrophysics, Mumbai, p. 129

Nalewajko, K., Giannios, D., Begelman, M. C., Uzdensky, D. A., & Sikora, M. 2011, Mon. Not. Roy. Astr. Soc., 413, 333

Ohsaki S., Shatashvili N. L., Yoshida Z., Mahajan S. M., 2002, Astrophys. J., 570, 395

Øieroset, M., Phan, T. D., Eastwood, J. P., et al. 2011, Phys. Rev. Lett., 107, 165007

Oka, M., Fujimoto, M., Shinohara, I., & Phan, T. D. 2010a, J. Geophys. Res., 115, A08223

Oka, M., Phan, T., Krucker, S., Fujimoto, M., & Shinohara, I. 2010b, Astrophys. J., 714, 915

Ossendrijver, M. 2003, Ast. Astrophys. Rev, 11, 287

Parfrey, K., Beloborodov, A. M., & Hui, L. 2012, Astrophys. J., 754, L12

Parfrey, K. P., & Menou, K. 2007, Astrophys. J., 667, L207

Parker, E. N. 1955, Astrophys. J., 122, 293

Paxton, B., Bildsten, L., Dotter, A., et al. 2011, Astrophys. J. Suppl., 192, 3

Martínez Pillet, V. 2013, Space Sci. Rev., 21

Piran, T. 2004, Rev. Mod. Phys., 76, 1143

Pitts, E., & Tayler, R. J. 1985, Mon. Not. Roy. Astr. Soc., 216, 139

Priest, E. & Forbes, T. 2000, Magnetic Reconnection (Cambridge University Press)

Pritchett, P. L. 2008, Phys. Plasmas, 15, 102105

Roytershteyn, V. & Daughton, W. 2008, Phys. Plasmas, 15, 082901

Schindler, K., Hesse, M., & Birn, J. 1988, J. Geophys. Res., 93, 5547

Schmitz, H. & Grauer, R. 2006, Phys. Plasmas, 13, 092309

Schwenn R., et al., 2006, Space Sci. Rev., 123, 127

Shibata, K. & Tanuma, S. 2001, Earth Planets Space, 53, 473

Sironi, L. & Spitkovsky, A. 2011, Astrophys. J., 741, 39

Speiser, T. W. 1965, J. Geophys. Res., 70, 4219

Spitkovsky, A. 2008, Astrophys. J., 673, L39

Spitzer, L. Jr. 2006, The Physics of Fully Ionized Gases, New York: Dover

Spruit, H. C. 1999, Astron. & Astrophys., 349, 189

Spruit, H. C. 2002, Astron. & Astrophys., 381, 923

Takahashi, H. R., Kudoh, T., Masada, Y., & Matsumoto, J. 2011, Astrophys. J., 739, L53

Tanaka, K. G., Fujimoto, M., Badman, S. V., & Shinohara, I. 2011, Phys. Plasmas, 18, 022903

Tassoul, J.-L. 1978, Princeton Series in physics, Princeton: University Press

Sturrock P. A., 1991, Astrophys. J., 380, 655

Thompson, C. & Duncan, R. C. 1995, Mon. Not. Roy. Astr. Soc., 275, 255

Woltjer L., 1958, Proc. Natl. Acad. Sci. USA, 44, 489

Uzdensky, D. A., Cerutti, B., & Begelman, M. C. 2011, Astrophys. J., 737, L40

Uzdensky, D. A., Loureiro, N. F., & Schekochihin, A. A. 2010, Phys. Rev. Lett., 105, 235002

Vasyliunas, V. M. 1975, Rev. Geophys. Space Phys., 13, 303

Velikhov, E. P. 1959, J. Exp. Theoret. Phys. (USSR), 36, 1398

Volponi, F., Yoshida, Z., & Tatsuno, T. 2000, Phys. Plasmas, 7, 2314

Vishniac, E. T. 2009, Astrophys. J., 696, 1021

Von Zeipel, H. 1924, in Probleme der Astronomie, Festschrift für H. v. Seeliger, ed H. Kienle, (Springer, Berlin), p. 144

Widrow, L. M., Ryu, D., Schleicher, D. R. G., et al. 2012, Space Sci. Rev., 166, 37

Workman, J. C., & Armitage, P. J. 2008, Astrophys. J., 685, 406

Yamada, M., Kulsrud, R., & Ji, H. 2010, Rev. Mod. Phys., 82, 603

Yin, L., Daughton, W., Karimabadi, H., et al. 2008, Phys. Rev. Lett., 101, 125001

Zenitani, S. & Hesse, M. 2008, Astrophys. J., 684, 1477

Zenitani, S. & Hoshino, M. 2001, Astrophys. J., 562, L63

Zenitani, S. & Hoshino, M. 2005, Phys. Rev. Lett., 95, 095001

Zenitani, S. & Hoshino, M. 2007, Astrophys. J., 670, 702

Zenitani, S. & Hoshino, M. 2008, Astrophys. J., 677, 530

Zhang, B. & Yan, H. 2011, Astrophys. J., 726, 90

Zhang J., Dere K. P., Howard R. A., Kundu M. R., White S. M., 2001, Astrophys. J., 559, 452

Vita

Daniel Ross Kagan was born in Houston, Texas in 1982, the son of Dr. Martin H. Kagan and Dr. Deborah L. Kagan. He received a Bachelor of Science degree in Physics in 2005 from Harvey Mudd College. He entered graduate studies in the Department of Astronomy at the University of Texas at Austin in 2005, and received an M.A. in Astronomy in August 2007. He took a leave of absence to participate in a research internship at the Weizmann Institute for Science in Rehovot, Israel from September 2007-February 2008, and reentered the graduate program in the Department of Astronomy at the University of Texas at Austin at a doctoral level at the University of Texas at Austin in September 2008.

Permanent address: 8907 Birdwood Court
Houston, Texas 77096

This dissertation was typeset with LaTeX[†] by the author.

[†] LaTeX is a document preparation system developed by Leslie Lamport as a special version of Donald Knuth's TeX Program.